

181 GGTCTCTTCCACCTGCTCTACACTGCTGCTTTTGTGTGACTTATCAAAATGGCC 240
181 GGTCTCTTCCACCTGCTCTACACTGCTGCTTTTGTGTGACTTATCAAAATGGCC 240
241 GGGATGAGAGATTTCTTCAATCCGAGTAATCAAACTTATTTCTTTAGAA 300
241 GGGATGAGAGATTTCTTCAATCCGAGTAATCAAACTTATTTTCTTTAGAA 300
301 CACAGTATTTCTGATCTTCTCATGATTCGACTTTTCCATTCAAAATCTTGTAGTGC 360
301 CACAGTATTTCTGATCTTCTCATGATTCGACTTTTCCATTCAAAATCTTGTAGTGC 360
301 CACAGTATTTCTGATCTTCTCATGATTCGACTTTTCCATTCAAAATCTTGTAGTGC 360
361 CAACCTGGGAACGACGACCTGGAACCTTTGTGTGCAAGTACCTCCCTCATATT 420
361 CAACCTGGGAACGACGACCTGGAACCTTTGTGTGCAAGTACCTCCCTCATATT 420
421 TTTTCAAGTATATAGATATTTTCTTCCGAGGATGATTAATGATGATGATGATGAT 480
421 TTTTCAAGTATATAGATATTTTCTTCCGAGGATGATTAATGATGATGATGATGAT 480
481 GACACACGAGCCATTTTAAACATCAACCCAAAATCTTGTGGGCTTAAGATTCCT 540
481 GACACACGAGCCATTTTAAACATCAACCCAAAATCTTGTGGGCTTAAGATTCCT 540
541 TGTTCATCTGGGCACTCATGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 600
541 TGTTCATCTGGGCACTCATGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 600
601 GCAGCGAGAGACAAAGTGAAGAAATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 660
601 GCAGCGAGAGACAAAGTGAAGAAATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 660
661 CTGGCATGAATAGTAAATATCATCTCTCAAGTCAATTTCTGATTAATTTCTTAAT 720
661 CTGGCATGAATAGTAAATATCATCTCTCAAGTCAATTTCTGATTAATTTCTTAAT 720
721 TATTTGATGTTATACATCTATTCAAAAGACGTCACGCTCATAGTAAGAACGAGGG 780
721 TATTTGATGTTATACATCTATTCAAAAGACGTCACGCTCATAGTAAGAACGAGGG 780
781 TGTAGTAAAGTCCGAGGAAAGGTGAAGCTCAAGTCTTCTTCTTCTTCTTCTTCT 840
781 TGTAGTAAAGTCCGAGGAAAGGTGAAGCTCAAGTCTTCTTCTTCTTCTTCTTCT 840
841 CTTTATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 900
841 CTTTATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 900
901 GGTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 960
901 GGTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 960
961 GTTAACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1020
961 GTTAACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1020
1021 CAGAAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1080
1021 CAGAAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1080
1081 CATATGAGAAAGAAAGACAGATGCTGACCCAAATGAAGACATCCATGTAAACAA 1140
1081 CATATGAGAAAGAAAGACAGATGCTGACCCAAATGAAGACATCCATGTAAACAA 1140
1141 TTTAACTAAAGAAATATTTCAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1187
1141 TTTAACTAAAGAAATATTTCAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1187

RESULT 2
US-09-827-937A-1
; Sequence 1, Application us/09827937A

Patent No. US20020052043A1
GENERAL INFORMATION:
APPLICANT: LI, YI
APPLICANT: Ruben, Steven M.
TITLE OF INVENTION: Human G-Protein Coupled Receptors
FILE REFERENCE: 1488,1220003
CURRENT APPLICATION NUMBER: US/09/827,937A
PRIOR FILING DATE: 2001-04-09
PRIOR APPLICATION NUMBER: 08/852,824
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 2247
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (226)..(1251)
US-09-827-937A-1

Query Match 97.7%; Score 1160; DB 10; Length 2247;
Best Local Similarity 100.0%; Pred. No. 1,8e-290;
Matches 1160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

28 CTGACAGCTGCAATTAATCTACTACTGATATCATTTCAAACTCCAGATCAACAGTT 87
147 CTGACAGCTGCAATTAATCTACTACTGATATCATTTCAAACTCCAGATCAACAGTT 206
88 ATAGATTAACCAACAAAGAAATGCAACCCGTCGACAAATTCACCTCTGCGCTGGAAAC 147
207 ATAGATTAACCAACAAAGAAATGCAACCCGTCGACAAATTCACCTCTGCGCTGGAAAC 266
148 CAGTCTGTGACAGAGAGACTACAAATCAACCCAGTCTCTTCCACTGCTACACTGT 207
267 CAGTCTGTGACAGAGAGACTACAAATCAACCCAGTCTCTTCCACTGCTACACTGT 326
208 CCTGTTTTTGTGGAATATCAAAATGGCTGCGATGAGATTTCTTCAAAATCCG 267
327 CCTGTTTTTGTGGAATATCAAAATGGCTGCGATGAGATTTCTTCAAAATCCG 386
268 GAGTAATCAAACTTATTTATTTCTTAAAGACAGATCTTCTGATCTTCTCATGTAT 327
387 GAGTAATCAAACTTATTTATTTCTTAAAGACAGATCTTCTGATCTTCTCATGTAT 446
328 TCTGACTTTTCCATTCAAAATCTTATGATGCAAACTGGGAACGACCACTGGAAC 387
447 TCTGACTTTTCCATTCAAAATCTTATGATGCAAACTGGGAACGACCACTGGAAC 506
388 TTTTGTGTGCAAGTAACTCCGCTCATTTTATTTTCAATGTATATGATATTTCTTCT 447
507 TTTTGTGTGCAAGTAACTCCGCTCATTTTATTTTCAATGTATATGATATTTCTTCT 566
448 CCGGAGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 507
567 CCGGAGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 626
508 CCGCAAAATCTTCTTGGGGCTTAAGATCTCTCTTGTTCATCTGGGCAATTCATGTTCT 567
627 CCGCAAAATCTTCTTGGGGCTTAAGATCTCTCTTGTTCATCTGGGCAATTCATGTTCT 686
568 ACTCTTTTGGCTTAACATATCTGACCAACAGGAGCGGAGAGACAAAGTGAAGAA 627
687 ACTCTTTTGGCTTAACATATCTGACCAACAGGAGCGGAGAGACAAAGTGAAGAA 746
628 ATGCTTTTCCCTTAATTCAGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 687
747 ATGCTTTTCCCTTAATTCAGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 806
688 TCAAGTCAATTTCTGATTAATTTCTTAATGTTATGTTATGTTATGTTATGTTATGTTAT 747
807 TCAAGTCAATTTCTGATTAATTTCTTAATGTTATGTTATGTTATGTTATGTTATGTTAT 866


```

: TITLE OF INVENTION: 15625 Receptor, A No. US20020156246A1e1 G-Protein Coupled Receptd
: FILE REFERENCE: 5800-13, 035800-171548
: CURRENT APPLICATION NUMBER: US/09/964, 008
: CURRENT FILING DATE: 2001-09-26
: PRIOR APPLICATION NUMBER: 09/187,134
: PRIOR FILING DATE: 1998-11-06
: NUMBER OF SEQ ID NOS: 5
: SOFTWARE: Patentln Ver. 2.0
: SEQ ID NO 4
: LENGTH: 2272
: TYPE: DNA
: ORGANISM: Macaca sp.
: US-09-964-008-4

Query Match      93.4%; Score 1108.8; DB 9; Length 2272;
Best Local Similarity 97.2%; Pred. No. 3.1e-277;
Matches 1128; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 28 CTCGAGCGTCGAACTACTACTACTGATACATTCGAACCCCTCCAGATCAACAGT 87
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 160 CTCGAGCGTCGAACTACTACTACTGATACATTCGAACCCCTCCAGATCAACAGT 219

QY 88 ATCAGGTAAACCAAGAAATGCAAGCCGTCGACAACTCAGCTCGCGCTGGGAAC 147
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 220 ATCAGGTAAACCAAGAAATGCAAGCCGTCGACAACTCAGCTCGCGCTGGGAAC 279

QY 148 CAGTCTGTGCACAGAGACTACAAATCACCAGGTCCTCTCCACTGCTACACTGT 207
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 280 CAGTCTGTGCACAGAGACTACAAATCACCAGGTCCTCTCCACTGCTACACTGT 339

QY 208 CCGTGTGTTGTTGAGCTTTCACAAATGGCCGGGAGTAGATTTTCTTCAATCCG 267
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 340 CCGTGTGTTGTTGAGCTTTCACAAATGGCCGGGAGTAGATTTTCTTCAATCCG 399

QY 268 GAGTAATCAAACTTTATTTTCTTAAGAACACAGTCTTGATCTTCATGAT 327
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 400 GAGTAATCAAACTTTATTTTCTTAAGAACACAGTCTTGATCTTCATGAT 459

QY 328 TCTGACTTTTCATTCAAAATTTCTTAGTGATGCCAAACTGGGAACAGGACCTGAGAC 387
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 460 TCTGACTTTTCATTCAAAATTTCTTAGTGATGCCAAACTGGGAACAGGACCTGAGAC 519

QY 388 TTTTGTGTGTCAGTACCTCCGTCATTTTCTTCAAGATGATATATCATGTTTCAT 447
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 520 TTTTGTGTGTCAGTACCTCCGTCATTTTCTTCAAGATGATATATCATGTTTCAT 579

QY 448 CCTGGAGCTGATACATGATGCTACACAGAAAGCCAGGCACTTTAAACATCCAA 507
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 580 CCTGGAGCTGATACATGATGCTACACAGAAAGCCAGGCACTTTAAACATCCAA 639

QY 508 CCCCCAAAATCTCTGGGGGCTAAGATCTCTGTGTGTCATCTGGGCATTCATGTTCT 567
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 640 CCCCCAAAATCTCTGGGGGCTAAGATCTCTGTGTGTCATCTGGGCATTCATGTTCT 699

QY 568 ACTCTCTTGGCCTAATCATGATTTCTGACCAACAGGACCGAGAGACAAGAAATGGAAGA 627
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 700 ACTCTCTTGGCCTAATCATGATTTCTGACCTAAGCGGCCCAAGAGACAAGATGGAAGA 759

QY 628 ATGCTCTTCTTAAATTCAGAGTTGGGTAGTCTGTCGATGGAATAGTAATTCATCTG 687
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 760 ATGCTCTTCTTAAATTCAGAGTTGGGTAGTCTGTCGATGGAATAGTAATTCATCTG 819

QY 688 TCAAGCATTTTCTGATTAATTTCTTAATTTGATTTGATTTGATTAACACTCACTTCAAA 747
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 820 TCAAGCATTTTCTGATTAATTTCTTAATTTGATTTGATTTGATTAACACTCACTTCAAA 879

QY 748 AGAAGCTACCGGTCATACGTAGTAAGAGAGGGGTAGGAAAGTCCCAAGAAAGG 867
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 880 AGAAGCTACCGGTCATACGTAGTAAGAGAGGGGTAGGAAAGTCCCAAGAAAGG 939

QY 808 GAAAGCTCAAAAGTTTCATTAATCTGCTATCTTATTTGTTGTTGTTGTTGTTGTTGTT 867
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 940 GAAAGCTCAAAAGTTTCATTAATCTGCTATCTTATTTGTTGTTGTTGTTGTTGTTGTT 999

```

```

QY 868 TGCCCGAATTCCTTACACCCCTGAGCCAAACCCGGGATGCTTTGACTGCACACTGCGAAA 927
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1000 TGCCCGAATTCCTTATACCTGAGCCAAACCCGGGATGCTTTGACTGCACACTGCGAAA 1059

QY 928 TACTCTGTTTATGTAAGAGAGACTGCTGTGTTAACTTCTTAAATGATGCTCGA 987
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1060 TACTCTGTTTATGTAAGAGAGACTGCTGTGTTAACTTCTTAAATGATGCTCGA 1119

QY 988 TCCGTTGATCATTTTCTTCTTGGCAAGTCTCAGAAATTCCTGATGATGCTGAA 1047
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1120 TCCGTTGATCATTTTCTTCTTGGCAAGTCTCAGAAATTCCTGATGATGCTGAA 1179

QY 1048 GTGCCCAATTCGCAACATCTCTGTCCAGAGCAATAGAAAAAAGAACAGATGCTG 1107
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1180 GTGCCCAATTCGCAACATCTCAGTCCAGAGCAATAGAAAAAAGAACAGATGCTG 1239

QY 1108 TGACCCAAATGAAGAGAGCTCCATGTAACAAATTAAGTAAGTAATTTCAATCTCT 1167
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1240 TGACCCAAATGAAGAGAGCTCCATGTAACATATTAAGTAAGTAATTTCAATCTCT 1299

QY 1168 TGTGTCAGAACTGCTTAA 1187
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1300 TGTGTCAGAACTGCTTAA 1319

RESULT 5
US-09-780-576-1
: Sequence 1, Application US/09780576
: Patent No. US20020072072A1
: GENERAL INFORMATION:
: APPLICANT: Civeili, Olivier
: APPLICANT: No. US20020072072A1hacker, Hans-Peter
: APPLICANT: Wang, Zhilwei
: APPLICANT: Relnscheld, Rainer
: TITLE OF INVENTION: Adp-Glucose Receptor
: FILE REFERENCE: P-UC 4530
: CURRENT APPLICATION NUMBER: US/09/780,576
: CURRENT FILING DATE: 2001-02-09
: PRIOR APPLICATION NUMBER: US 60/234,025
: PRIOR FILING DATE: 2000-09-20
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: fastSeq for Windows Version 4.0
: SEQ ID NO 1
: LENGTH: 1143
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (32)...(1077)
: US-09-780-576-1

Query Match      92.1%; Score 1093.8; DB 10; Length 1143;
Best Local Similarity 99.8%; Pred. No. 1.8e-273;
Matches 1095; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 91 AGGTAACAACAAGAAATGCAAGCCGTCGACATCTACCTGCGCTGGGAACACAG 150
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 36 AGGTAACAACAAGAAATGCAAGCCGTCGACATCTACCTGCGCTGGGAACACAG 95

QY 151 TCTGTGCACAGAGACTACAAAATCACCAAGTCTCTTCCACTGCTCTACACTGTCT 210
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 96 TCTGTGCACAGAGACTACAAAATCACCAAGTCTCTTCCACTGCTCTACACTGTCT 155

QY 211 GTTTTGTGTCGATATACAAATGGCCGGGAGATGAGATTTCTTCAATCCGGAG 270
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 156 GTTTTGTGTCGATATACAAATGGCCGGGAGATGAGATTTCTTCAATCCGGAG 215

QY 271 TAAATCAACTTATATTTTCTTAAAGAACAGAGTCTTGATCTCTCATGATTC 330
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 216 TAAATCAACTTATATTTTCTTAAAGAACAGAGTCTTGATCTCTCATGATTC 275

QY 331 GACTTTTCATTCAAAATTTCTTAGTGATGCCAAAATCTGGGAACAGACACTGAGAACTTT 390

```



```

|||||
Db 276 GACTTTCATCAAAATTTCTAGTATGCCAACTGGGAACAGACACAGCACTG 335
QY 391 TGTGTCAAGTACCTCCGATATTTTATTTTCAATGTATATCATTCATTCCT 450
Db 336 TGTGTCAAGTACCTCCGATATTTTATTTTCAATGTATATCATTCATTCCT 395
QY 451 GGGACGTATACATATCATGCTGCTACAGAAAGACCAGCCAGCCATTTAAACATCCACC 510
Db 396 GGGACGTATACATATCATGCTGCTACAGAAAGACCAGCCAGCCATTTAAACATCCACC 455
QY 511 CAAAAATCTCTGGGGCTAAGATTTCTCTGTTGTCATCTGGGCATTCATGTTCTACT 570
Db 456 CAAAAATCTCTGGGGCTAAGATTTCTCTGTTGTCATCTGGGCATTCATGTTCTACT 515
QY 571 CTCTTTCCTTAATCAGATTCGACCAAGACCCGAGAGACAAGATGGAAGAAATG 630
Db 516 CTCTTTCCTTAATCAGATTCGACCAAGACCCGAGAGACAAGATGGAAGAAATG 575
QY 631 CTCTTTCCTTAATCAGATTCGACCAAGATTCGACCAAGATGGAAGAAATG 690
Db 576 CTCTTTCCTTAATCAGATTCGACCAAGATTCGACCAAGATGGAAGAAATG 635
QY 691 AGTCAATTTTGGATTAATTTCTTAATTTGTTATGTTATACATCATTAACAAAGA 750
Db 636 AGTCAATTTTGGATTAATTTCTTAATTTGTTATGTTATACATCATTAACAAAGA 695
QY 751 ACTGTACCGGTCATACAGTAAGAAAGAGGGGTAGTAAAGTCCCGAGAAAGGTGAA 810
Db 696 ACTGTACCGGTCATACAGTAAGAAAGAGGGGTAGTAAAGTCCCGAGAAAGGTGAA 755
QY 811 CGTCAAGATTTTCAATATCATGCTGATATTTCTTAATTTGTTTCTCTTCCATTTTGC 870
Db 756 CGTCAAGATTTTCAATATCATGCTGATATTTCTTAATTTGTTTCTCTTCCATTTTGC 815
QY 871 CCGAATTCCTTACACCTGAGCCAAACCCGGGATGTTCTTACGTCGCTGTGAATATC 930
Db 816 CCGAATTCCTTACACCTGAGCCAAACCCGGGATGTTCTTACGTCGCTGTGAATATC 875
QY 931 TCTGTCTATGTGAAGAGACACTGTGTGTTAACTTCTTAATGATCCCTGAGATCC 990
Db 876 TCTGTCTATGTGAAGAGACACTGTGTGTTAACTTCTTAATGATCCCTGAGATCC 935
QY 991 GTTCATCTATTTTCTTCTTGGCAAGTCTCTGCAAAATCTCTGATATGCTGAAGTG 1050
Db 936 GTTCATCTATTTTCTTCTTGGCAAGTCTCTGCAAAATCTCTGATATGCTGAAGTG 995
QY 1051 CCCCATTCTGCAACATCTCTGTCGCCAGAACATAGAGAAAAAGAGATGCTGGA 1110
Db 996 CCCCATTCTGCAACATCTCTGTCGCCAGAACATAGAGAAAAAGAGATGCTGGA 1055
QY 1111 CCCCATTCTGCAACATCTCTGTCGCCAGAACATAGAGAAAAAGAGATGCTGGA 1170
Db 1056 CCCCATTCTGCAACATCTCTGTCGCCAGAACATAGAGAAAAAGAGATGCTGGA 1115
QY 1171 GTTCAGAACTGTTAAA 1187
Db 1116 GTTCAGAACTGTTAAA 1132

```

```

; CURRENT APPLICATION NUMBER: US/09/769,159
; CURRENT FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 08/956,975
; PRIOR FILING DATE: 1997-10-23
; PRIOR APPLICATION NUMBER: 09/221,456
; PRIOR FILING DATE: 1998-12-28
; PRIOR APPLICATION NUMBER: 09/558,740
; PRIOR FILING DATE: 2000-04-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1124
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
; US-09-769-159-1

Query Match      25.7%; Score 305.6; DB 10; Length 1124;
Best Local Similarity 57.7%; Pred. No. 1,4e-69;
Matches 545; Conservative 0; Mismatches 399; Indels 0; Gaps 0;

QY 152 CTGTGACAGAGACTACAAATATCACCAGGTCTCTCCGACTGCTCTACATCTGCTG 211
Db 137 CGGTGCCCCAGAGACACTCGATATGATACAGTGTATTCAGCCCTCTACACAGTGT 196
QY 212 TTTTGTGGACTTATCACAATATGCGCTGCGATGAGATTTTCTTCAATCCGAGAT 271
Db 197 TTTTGGACCGCATCTGCTGATATCTTGGCTCTGAGGGTGTGTTGATCATATCCAGC 256
QY 272 AAATCAACTTATATTTTCTTAAGACAGATCATTTGTATCTTCTCATATCTG 331
Db 257 TCTTCACCTTCATCATCTTACCTCAAAACACTTGGTGGCCACTGTATATATGACATC 316
QY 332 ACTTTCATCAAAATTTCTAGTATGATGCCAACTGGGAACAGAGACCAGTGAATCTTT 391
Db 317 ATGCTTCCTTCAAAATCTCTGACTACACCTGCGACCTGCGACCTGAGCTTTT 376
QY 392 GTGTGCAAGTTACCTCGCTCATATTTTATTTCAATGTATATCATATTTCTCTG 451
Db 377 GTGTGCTGTTTCTTCTGCTGATATTTATGACACATGTATGTGGCATCTGCTGTA 436
QY 452 GGAATGATATCATGATCTGCTACAGAGACACAGCCATTTAAACATCCACCC 511
Db 437 GGGCTCANTAGCCTTGCAGAGATCTCTCAAGATCATACACCTTTGAGAAATATTTCTA 496
QY 512 AAAATCTCTTGGGGCTAAGATCTCTGTTGATCATCTGGGATCATCTTCTACTC 571
Db 497 AAAAATCTGTTTGGAAAAAGCTCATCTTCAATGCTGTTTCTTCTTCTATC 556
QY 572 TCTTGGCTTAACATGATTTGACCAACAGCAGCCGAGAGACAAGATGGAAGAAATGC 631
Db 557 TCCCTGCCAAATAGCATCTTGGACACAAAGAACCAACACCATCTGCTGTAAGAAAGTG 616
QY 632 TCTTCTTAATACAGATTTGGTCTAGTCTGGGATGAATAGTAATTAATCTGTCAA 691
Db 617 GCTTCTTAAGGGGCTCTGGGGCTGAATGGATCAAAAGTAATTAATATATGTCAG 676
QY 692 GTCAATTTCTGATTAATTTCTTAATTTGTTATTTATTTATACACTATACAAAGAA 751
Db 677 TTTATTTCTGACTGTTTATATCTAATGCTTGTGTTTATGTGTTATGTGCAAAAAA 736
QY 752 CTGTACCGGTATACGTAGAACAGAGGGGTAGATAAGTCCCGAGAAAAAGGTGAC 811
Db 737 GTATATGATTTCTTATAGAAATCTCCAAAGGTAGAACACAAAAAACAACAAAGCTGGA 796
QY 812 GTCAAGTTTCTATATCATATGCTGTAATCTTATTTGTTTCTTCTTCCATATTTGCC 871
Db 797 GGCAGAGATTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 856
QY 872 CGAATTCCTTACACCTGAGCCAAACCCGGGATGCTTTGATCTGACATGCTGAATATCT 931
Db 857 AGATTTTCATATCTACAGTCAACCAACAAATAGAGTGTAGTGTAGTGTAGTGTAGTGTAG 916
QY 932 CTGTCTATGTGAAGAGACATCTGTGTGTTAACTTCTTAAATGATGCTGATGCTG 991

```

```

RESULT 6
; Sequence 1, Application US/09769159
; Patent No. US20010021509A1
; GENERAL INFORMATION:
; APPLICANT: SATHE, GANESH M.
; APPLICANT: CHAMBERS, WENDY S.
; APPLICANT: CHAMBERS, JON
; APPLICANT: MOIR, ALISON
; APPLICANT: SEKERES, PHILIP
; TITLE OF INVENTION: METHODS OF SCREENING FOR AGONISTS AND
; TITLE OF INVENTION: ANTAGONISTS OF THE HNEA81 RECEPTOR
; FILE REFERENCE: GH-70318-DB3

```

```

Db 917 CTTTATTTGCTAAGAAACACTCTCTTTTGGCAGCAACTAATTTGTATGATCCC 976
Qy 992 TTTCATCTATTTTCCCTTTGCAAGTCCCTCAGAAATTCCTGTATGATGCTGAAGTGC 1051
Db 977 TTTAATATACATATTTCTATGTATAAAAAATTCCAGAAAAAGGTACCATGTATGCAAGGAGA 1036
Qy 1052 CCCAATTCGCAACATCTCTGTCCCGAGCAATAGCAAAAAAGA 1095
Db 1037 AAGACACAGACATCAAGCAAGAAATCATAGCAGTCAGACAGA 1080

```

RESULT 7

```

US-09-848-889-2
; Sequence 2, Application US/09848889
; Patent No. US2002002555A1
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Guegler, Karl J.
; APPLICANT: Cheng, Muzong
; TITLE OF INVENTION: GPCR DIAGNOSTIC FOR BRAIN CANCER
; FILE REFERENCE: PC-0042 CIP
; CURRENT APPLICATION NUMBER: US-09/848,889
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 1488
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US2002002555A1 568987CBI
US-09-848-889-2

```

```

Query Match 25.5%; Score 302.4; DB 10; Length 1488;
Best Local Similarity 57.5%; Pred. No. 1e-68; Indels 0; Gaps 0;
Matches 543; Conservative 0; Mismatches 401;

```

```

Qy 152 CTTGTCAGCAGAGACTACCAAAATCACCCAGTCTCTCCCACTGCTACACTGTCTG 211
Db 131 CGGTGCGCCAGAGACACTGGAGTAGACAGCTGTATTCGCCAGCCCTACACAGTGTT 190
Qy 212 TTTTGTGTGACTTATTCACAAATGCGCTGGCGATGAGATTTTCTTCAATCCGAGT 271
Db 191 TTTCTGACCGGATCTCTGTGTAATCTTGGCTGTGGTGTGTTCACATCCCACG 250
Qy 272 AATCAAACTTTATTTTCTTGAAGACACAGTCATTTTGATGCTGCATGATTCG 331
Db 251 TCTCCACCTTCATCTACCTCAACAAACACTTTGTGGCGCATGTGAATGACACTC 310
Qy 332 ACTTTTCATCAAAATTTCTTAGTAGTGCACAACTGGGAACAGACACACTGAGAACTTT 391
Db 311 ATGCTTCCTTTCAAAATTCCTGCTGACCTGACACCTGGCACTGGAGCTCAGAGCTTT 370
Qy 392 GTGTGCAAGTACCTCCCTCATATTTTATTTTACAAATGATATATGATTTTCTTCG 451
Db 371 GTGTGCGTGTCTTCTCGGTGATATTTTATGAGACCATGATGTGGGCGTGTGTTA 430
Qy 452 GGAATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 511
Db 431 GGGCTATATGCTTTGACAGATTTCCCTCAAGATCATGAGACTTTGGAATATTTTCTTA 490
Qy 512 AAAATCTCTTGGGGCTAAGATTTCTCTGTGTGATCTGGGCATTTCTGTTCTACTC 571
Db 491 AAAAATCGTGTGTTGCAAAAAGGCTCATCTTGTGTGTTTCTTCTTCTTCTC 550
Qy 572 TCTTTGCTCAATGATTTGTGACCAAGCGACGACGACGACGACGACGACGACGACG 631
Db 551 TCCCTGCAATATGATCTTGTGAGCAACGAGCAACCACTGCTGTGAAAAAGTGT 610
Qy 632 TCTTTCTTAAATCAAGAGTGTGCTGATGCTGCAATGAATGATATGATCTCTCA 691

```

```

Db 611 GCTTCTTAAAGGGCCCTCGGGGCTGAATGGCATCAATGGTAAATACATATGCCAG 670
Qy 692 GTCATTTCTGATTAATTTCTTAATTTGATATGATATGATACACTATTCACAAAAGA 751
Db 671 TTTATTTCTGAGCTGTTTAACTATATGCTGTGTTTATGCTTATGCAAAAAA 730
Qy 752 CTGTACCGGTATCTAAGAACGAGGGGTATGATTAAGTCCCGAGAAAAAGTGAAC 811
Db 731 GTATATGATCTTATGAAAGTCCAAATGTAAAGACAGAAAAACAACAAAAGCTGGA 790
Qy 812 GTCAAGTTCTATCTATCTATGCTGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 871
Db 791 GGCAGATATTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 850
Qy 872 CGAATTCCTTACACCTGAGCCAAACCCGGAGTGTGTTGACTGCTGCTGAAATATCT 931
Db 851 AAGAGTTCATATCTACAGTCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 910
Qy 932 CTGTCTATGTGAAGAGACACTGTGTGTTACTCTTAAATGATGATGCTGATGCC 991
Db 911 CTGTATTTGCTAAGAAACACTCTCTTTTGGCAGCACTAACATTTGTATGATGCC 970
Qy 992 TTTCATATTTTCCCTTGAAGTCTCCAGAAATTCCTGTATGATGATGCTGAAGTGC 1051
Db 971 TTTATTCATATTTCTTATGTAAAAATTCACGAAAGCTCCATCATGTATGCAAGGAGA 1030
Qy 1052 CCCAATTCGCAACATCTCTGTCCAGACAAATAGCAAAAAAGA 1095
Db 1031 AAGACACAGACATCAAGCAAGAAATCATAGCAGTCAGACAGA 1074

```

RESULT 8

```

US-09-919-497-27
; Sequence 27, Application US/09919497
; Patent No. US2002010662A1
; GENERAL INFORMATION:
; APPLICANT: Mutter, George L.
; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
; FILE REFERENCE: B0801/7225
; CURRENT APPLICATION NUMBER: US-09/919,497
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/221,735
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 27
; LENGTH: 2416
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-497-27

```

```

Query Match 21.3%; Score 252.4; DB 10; Length 2416;
Best Local Similarity 55.5%; Pred. No. 1.e-55;
Matches 487; Conservative 0; Mismatches 391; Indels 0; Gaps 0;

```

```

Qy 155 TGCACGAGACTACCAAAATCACCCAGTCTCTTCCACTGCTCTACACTGCTCTTT 214
Db 256 TCTCTCAGAACCTCTGATCACTGACAGATCATCTGCTGTGATGATGATGCTTC 315
Qy 215 TTTGTGAGCTTATCAAAATGCGCTGGGATGAGATTTTCTTCAATCCGAGTAAA 274
Db 316 ATTTGGGGAATCTTACTCATGATGATGATGATGATGATGATGATGATGATGATGAT 375
Qy 275 TCAAACTTATTTATTTTGAAGAACAGATCTTGTGATCTCTGATGATGATGATGAT 334
Db 376 AAGAGTTTATCTATCTATCTCAAGAAATTTGTTATGCTGATGATGATGATGATGAT 435
Qy 335 TTTTCATCAAAATTTCTAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 394
Db 436 TTTCTTCAAGATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 495
Qy 395 TGTCAAGTATCTCGTCTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 454

```

Db	100	ATCAGGGGGCTCCCTCCAAATGGGATPACAGAGATGATATCTTTATGTGGCCAGCTCC	159
Qy	275	TCAAACTTATATATTTTCTTAGACACAGCTATTTCTGATCTTCATGATTTGACT	354
Db	160	AAGGATTTATACATCTATCTCAGAACATAGTGTGGCTGACTCTTCAGAGGGCCGACT	219
Qy	335	TTTCCATTCCAAATTCCTTAGGATGGCCAACTGGGAAACAGACACTTTTGTG	394
Db	220	TTCCCTTTCAAAGTCTGTGGATCGTCAAGGCTCGGCCCTGGCAGGGTAATGTGTTGTG	279
Qy	395	TGTCAAGTTTACCTCCGTCATATTTTATTTTCCATATGATATACATTTATCTTCGGGA	454
Db	280	TGCAGGGCTCTGCCGTCATCTTCTATGTAATATGTACGTACAGCATCGTGTTCTTGGG	359
Qy	455	CTGATTAACATATCATCGCTACAGAACACACAGGCCATTTTAAACATCCAAACCCCAA	514
Db	340	CTCATACGCTTTGACAGGATCTATAAATATGTGAAGCCCTTTCGACGTATATGTGTACG	399
Qy	515	AATCTCTTGGGGGCTAAGATTTCTCTCTGTGTATCTTGGGCATTCATGTTCTACTCTCT	574
Db	400	TCGGGAACCTATAGCAGAGCTCTTCTCTGTGTCTGTGGATGTCCATAGCTTCTCCCTGCT	459
Qy	575	TTGCTTAACATGATTTCTGACCAACAGCAGCCGAGAACAAAGAAATGTGAAGAAATGCTCT	634
Db	460	GTCCCAAACATCATCTCTBACAAACAGGCTCTCAAGACGTCACGAAGATATACGTGATG	519
Qy	635	TTCTCTTAATCAGAGTTCCGTCTATGTCTGGCATGAAATAGTAAATTCATCTGTCAAGTC	694
Db	520	GAGCTCAAAAACAGAGCTGGGGCGGAAGTGGCAACAAAGCCGTCTAACATATCTTCTGTAGT	579
Qy	695	ATTTCCTGTGATTAATTTCTTAATTTGTATTTGTATATACACTCATTTACAAAGAATG	754
Db	580	ATCTCTCTGGGCTGATTTCTTCTCTATATGCTCTTCTACAGGCGCATCACAGAGAGATC	639
Qy	755	TACGGGTACATCATGTAAGAACAGGGGATGTAGTAAATGCCACAGAAAGATGGAACGTC	814
Db	640	TTCAAGTCTACCTCCACAGTCCAGGACAGAAATTCACCTCCGTCAAGAGAAAGTCCAGCGC	699
Qy	815	AAAGTTTCATTAATCATTTGCTATGCTATTTCTTTATTTGTTTGTCTTCCATTTTGGCCGA	874
Db	700	AATATCTTCACACATCTGTCTCTGTTTGTGTGTCTGTGCTTGTGCTTACACATTTGCCAGA	759
Qy	875	ATTCTTACACCTGTAGCCAAACCCGGGATGCTTTTACTGTGCACACTGCTGAATAATATCTG	934
Db	760	ATCCCTCACAAAGAGTCAAGCAGGAGGTCACTACAGCTGCGCGAGAGAGAACCTG	819
Qy	935	TTCTATGTGAAGAGAGACTCTGTGTATTTTAACTTCTTAATATGATCCCTGTGATCCGTTTC	994
Db	820	CTGTATGGAAAGAAATTCATCTCTCTACTCTCGGCTGCCAATGTGTGTGGAGCCCATTT	879
Qy	995	ATCATATTTTCTTGTGCAATGCTCTTACAGAAATCTTGAATAAATATGCT	1044
Db	880	ATTATATTTCTTATATGCGACCACTTTAAGAAATCTTTAAATAGAAGTT	929
RESULT 10			
US-09-826-508-9			
: Sequence 9, Application US/09826508			
: Patent No. US20010025099A1			
: GENERAL INFORMATION:			
: APPLICANT: Nabil Elshourbagy			
: APPLICANT: Lisa Vawter			
: TITLE OF INVENTION: G Protein-Coupled Receptor Polypeptides			
: TITLE OF INVENTION: and polynucleotides			
: FILE REFERENCE: GB-70744USB			
: CURRENT APPLICATION NUMBER: US/09/826,508			
: CURRENT FILING DATE: 2001-04-05			
: NUMBER OF SEQ ID NOS: 40			
: SOFTWARE: FastSeq for Windows Version 3.0			
: SEQ ID NO 9			
: LENGTH: 1077			
: TYPE: DNA			
: ORGANISM: HOMO SAPIENS			

US-09-826-508-9

Query Match	17.7%;	Score 209.8;	DB 10;	Length 1077;
Best Local Similarity	53.4%;	Pred. No. 8.1e-45;		
Matches 464;	Conservative 0;	Mismatches 402;	Indels 3;	Gaps 1.

QY	157	CACCAAGACATACAAATTCACACCGAGGCTCTTCCACCTGGCTCAGACGTCCGGTTTT	218
Db	102	CACCTTCACAAATTAATTTGACACAAATTTCTTCCCGTCTTTATCTCATTTATTTGTT	161
QY	217	TGTTGACCTATACCAAAATGGCCCTGGGAGTAGAGATTTTCTTTCAAAATCCGGAGTAATC	278
Db	162	GGCAAGACATCTTGCTGAATGGTTTAGAGAGTGTGATCTTCTCCACATTTAGGAATAAAC	221
QY	277	AAACTTATTTTCTTTCTTAAGAACAAGCATTTCTGATCTTCTCATGATTTGCACTTT	336
Db	222	CAGCTCATATTTATCTATCCAAAAACAATAGTGGTTGACAGACCTCATTAATGACGCTGAATT	281
QY	337	TCCATTCACAAATTTCTTAGTGATGCGCAACTGGGAAACAGAGACCACTGAGACTTTTGTG	396
Db	282	TCCATTTGGAATAGTCATGATGAGGATTTGGACCTTGGTACTTCAAGTTATTTCTCTG	341
QY	397	TCAAGTTACTCCGTCATATTTTATTTACAAATGATATAGTATTTTCATTCTCGGACT	456
Db	342	CAGATACACTTCAGTTTGTGTTTATGCAAACTATATACTTCATCTGTTCTTGGGCT	401
QY	457	GATTAATCATGATGATGCTACCGAAGACACACAGGCCATTTTAAAAATCCAAACCCCAAAA	516
Db	402	GATTAAGCATTTGATTCGCTATCTGAAAGTGTGTCMAGCCATTTGGGACTCTCGATGTACAG	461
QY	517	TCCTTTGGGGCTAAGATTCTCTGTGTTGCATCTGGGCAATCATGTTCTTACTCTCTT	576
Db	462	CATTAACCTTCAGAAAGTTTATCTGTTTGTGTTGGTATCATGAGCTGTTTGTGCTTT	521
QY	577	GCTACATCATGTTTGTGACCAACAGGACGGACGAGACAAAGAATGGAAGAAATGCTTT	636
Db	522	GCCAAACATATCTCTGCAAAATGCTCAGCCCAACAGAGCAATATTCATGACTGCTAAA	581
QY	637	CCCTAAATCAGAGTTGCGTCTAGCTGAGCATGAATAGTAAATTTACATCTGTCAGATCAT	696
Db	582	ACTTAAGAATCTTTGGGGGCAATATGGCATAGCGCAGTCACTATGTGAACAGTCTT	641
QY	697	TTTCTGATTAATTTCTTAATTTGTTATTTGTTATTTATACACTATTTCAAAACACTGTA	756
Db	642	GTTTGTGGCGCTGCTGATTTCTGATCGGATGTTAATATGACCATATCCAGGTATCCCA	701
QY	757	CCGGCTCATAGTAAGAACGAGGGGTAGTAAAGTCCCGAGAAAAGGTGAGACGCA	816
Db	702	CAATTCAGAGGCAATTCATTAAGT---CAGTCAAGCCGAAAGCAAAAATATACACAG	758
QY	817	AGTTTCAATATATCATGCTGATTTCTTTATTTGTTTGTTCCTTTCCATTTTCCCGAAT	876
Db	759	CATGAGGATGTGTGTGCTGTGTTTTTACTGCTGTTCTCCATATACACTTGTGCAAGAT	818
QY	877	TCTTACACCTGTGAGCCAAACCCGGGATGTCTTTGATGTCAGACGTGAATAATCTCTGT	936
Db	819	TCTTTTATCTTTTAGTCATCTTAGACAGGCTTTTAGATGAATGTGCACAAAAATCTATA	878
QY	937	CTATGTAAAGAGACACTGTGTGTTAACTTCTTAAATGCATGCTGTGATCCGTTCAT	996
Db	879	TTACTGCAAGAATAATTCACCTTTTCTGTGCGGTGATTAATGTTGCTGTGATCCAAAT	938
QY	997	CTATTTTCTCTTGCAAATGCTTCAGAA 1025	
Db	939	TTACTTTTATGTAGTGAATTTTCCAA 967	

```

RESULT 11
US-09-974-298-85
; Sequence 85, Application US/09974298
; Patent NO. US20020156263A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Huei-Mei

```

```

? TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
? FILE REFERENCE: PA-0037 P
? CURRENT APPLICATION NUMBER: US/09/974,298
? CURRENT FILING DATE: 2001-10-04
? PRIOR APPLICATION NUMBER: 60/238,331
? PRIOR FILING DATE: 2000-05-10
? NUMBER OF SEQ. ID NOS: 194
? SOFTWARE: PERL Program
? SEQ ID NO 85
? LENGTH: 1444
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: misc_feature
? OTHER INFORMATION: Incyte ID No. US20020156263A1 1650519CBH
US-09-974-298-85

```

Query Match	17.5%	Score 207.4;	DB 9;	Length 1444;
Best Local Similarity	53.0%;	Pred. No. 3.8e-44;		
Matches 467; Conservative	0;	Mismatches 411;	Indels 3;	Gaps 1

QY	157	CACAGAGACATACAAAATCCACCAGGCGCTCCACAGCGCTACAGCTGCTCGTTT	216
Db	420	CACCTTCACAAAGAAATTTACACAAATGTGCTTGGCGGCTTTATTCATATATTTGT	479
QY	217	TGTTGACCTATACAAATAGCGCTGGCGATGAGGATTTTCTTCAATCCGAGTAAATC	276
Db	480	GGCAAGCATCTTGCTGAATGTTTACGAGTGTGATCTTCCACATTTAGSAAATAAAC	539
QY	277	AAACTTATATTTTCTTAAAGACAGCTCATTTTGATCTTTCAGATGTTGCACTT	336
Db	540	CAGGTCATATTCCTACCTCAAAAACATAGTGGTGGACACTCATATATGCACTACAT	599
QY	337	TCCATCAAAATTTCTAGTATGCCAAACGCGGAACAGGACCACTGAGAACTTTGTG	396
Db	600	TCCATTTTCAATATGCTCATATGTCAGAGATTTGGACCTTGTACTTCAGATTTATCTCTG	659
QY	397	TCAAGTTACCTCGTCATATTTTATTTTCAAAATGATATACATATTCATCTCGGACT	456
Db	660	CAGATACACTTCAGTTTGTGTTTATGCAAAACATGATCTTCATCGTGTGCTGGCT	719
QY	457	GATACATATGATCGCTACCAAGAACCAACAGGCATTTAAACATCCAAACCCAAAA	516
Db	720	GATAGGATTGATCGCTATCTGAAAGGTGGTCAAGCCATTTGGGAGCTCGGATGTACG	779
QY	517	TCTCTGGGGGCTAAGATCTCTCTGTGTCATCTGGGCAATGTTCTTACCTCTT	576
Db	780	CATACCTTACCAAGGTTTATCTTCTTTGTGTTGGGATCAAGGCTGTTTCTCTT	839
QY	577	GCCTAACATGATTTCTGACCAACAGGACGGCGAGACAGAAATGTGAAGAAATGCTCTT	636
Db	840	GCCAAACATCATCTCTGCAAAATGTTGAGGCAACAGAGACATATCCATGACTGTCAA	899
QY	637	CCTTAAATCAGAGTGGGTAGCTGCGCATGAAATAGTAAATTAATCTGTCAAGCAT	696
Db	900	ACTTAAAGTCTTTGGGGGTCAAAATGCGATACGGGACACCTATGTGAACAGCTGCT	959
QY	697	TTTCTGATTAATTTCTTAATTTCTTATGTTATGTATGTTATACATCATTTACAAAAGACTGA	756
Db	960	GTTTGTGGCCGTCTCTGTGATTTCTGTATGGAATGTATACATAGCCATATCCAGGTATACCA	1019
QY	757	CCGGTCATACGTAAAGACGAGGGGTAGGTAAAGTCCCGAGAAAAGGTAAAGTCA	816
Db	1020	CAATCCAGAGGCATTCATTAAGT---CAGTCAGCCGCAAAAGCCAAACATMACAGAG	1076
QY	817	AGTTTCATTATCATGCTGATTTCTTATTTGTTTGTTCCTTCCATTTTGCCCGAT	876
Db	1077	CATCAGGGTGTGTGTGGCTGTGATTTTACGCTCTTACATATACACTGTGTGAGAT	1136
QY	877	TCTTACACCCCTGAGCCAAACCGGAGTGTCTTTGACTGCACTGCTGAAATACTCTGT	936
Db	1137	GCCCTTCACTTTAGTACATTAACAGAGGCTTTAGTATGATCTGACAAAATATCTCTTA	1196

Qy	937	CTATGTCGAAGAGCAGCTGTGTGTAATCTCTTAAAGCAGTCGATCCGTTCA	996
Db	1197	TTACTGCAGAAGAAATTCACATTTCTCTGTGTGCTGAATGTTGCCGTGATCCAAATAT	1255
Qy	997	CTATTTTCTCTGCAAGTCCTGCAGAAATTCCTTGATA	1037
Db	1257	TTACTTTTCTATGTGTAGTGATTTTCAGAAATGGCTGTGA	1297

```

RESULT 12
US-09-919-172-21
: Sequence 21, Application US/09919172
: Patent No. US20020119463A1
GENERAL INFORMATION:
APPLICANT: Farris, Mary
TITLE OF INVENTION: PROSTATE CANCER MARKERS
FILE REFERENCE: PA-0036 US
CURRENT APPLICATION NUMBER: US/09/919,172
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/222,469
PRIOR FILING DATE: 2000-07-28
NUMBER OF SEQ ID NOS: 102
SOFTWARE: PERL Program
SEQ ID NO 21
: LENGTH: 1444
: TYPE: DNA
: ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc-feature
OTHER INFORMATION: Incyte ID No. US20020119463A1 1650519CBI
US-09-919-172-21

```

Query Match	17.5%	Score 207.4;	DB 10;	Length 1444;
Best Local Similarity	53.0%;	Pred. No. 3.8e-44;		
Matches 467;	Conservative 0;	Mismatches 411;	Indels 3;	Gaps 1

OY	697	TTCGTGATTAATTCTTAAATTTCTAATGTAATACACTACATAAAAAGAACGTA	756
Db	960	GTTTGTCGCCCGCTGCTGGTAGTTCTGAATCCGATGTTTACAAGCAATATCAGATCATCCA	1019
OY	757	CCGGTCATACGTAAAGAACGAGGGGTGTAGGTAAAGTCCCCAGAAAAAGTGAACGTCAA	816
Db	1020	CAAAATCCAGACGAGCAATTCAATAGT---CAGTCAAGCCGMAAGCCMAAACATACACGAG	1076
OY	817	AGTTTTCAATATCATACGCCTGATCTCTTTATATTTGTTGGTCTCTTCATATTTGGCCGAT	876
Db	1077	CATGAGGTTGTGTGTGGCTGTGATTTTACGTCTTTCTACATATFACCTGTGCAGAT	1136
OY	877	TCCTTACACCCTGAGCCAAACCCGGAGTCTTTGACTGCACTGCTGAATAATCTCTGTT	936
Db	1137	GCCCTTCACTTTAGTCACTTACACGAGGCTTTAGATGTAATCTGCACAAAAAATCTCTATA	1196
OY	937	CTATGTAAAGAGACGACCTGTGTGTTAACTTCCTAAATCATGCATCCGATCCGTCAT	996
Db	1197	TTACTGCAAGAAATTTACACTTTTCTGTGCTGTGATATGTTGGCTGAGTCAATATAT	1256
OY	997	CTATTTTTCCTTTGCAAGTCCTTCAGAAATTCCTTGATAA	1037
Db	1257	TTACTTTTTCATGTGTAAGTCATTTTCAACAGATGCGCTGCA	1297

```

RESULT 13
US-09-848-889-9
: Sequence 9, Application US/09848889
: Patent No. US2002002555A1
: GENERAL INFORMATION:
: APPLICANT: Au-Young, Janice
: APPLICANT: Guegler, Karl J.
: APPLICANT: Cheng, Muzong
: TITLE OF INVENTION: GPCR DIAGNOSTIC FOR BRAIN CANCER
: FILE REFERENCE: PC-0042 CIP
: CURRENT APPLICATION NUMBER: US/09/848.889
: CURRENT FILING DATE: 2001-05-03
: NUMBER OF SEQ ID NOS: 12
: SOFTWARE: PERL Program
: SEQ ID NO 9
: LENGTH: 592
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: Incyte ID No. US2002002555A1 6867412H1
US-09-848-889-9

```

Query Match	16.68;	Score 197.4;	DB 10;	Length 592;
Best Local Similarity	59.38;	Pred. No. 9.9e-42;		
Matches 336;	Conservative 0;	Mismatches 231;	Indels 0;	Gaps 0;

Db 313 TGTATAGGCTATACGCTTTGACAGATTCCTCAAGATCATACACCTTTGAGAAATATTT 372
QY 507 ACCCAAAAATCTTGGGGCTAAGATTCTCTGTGTGATCTGGGCAATTCATCTTC 566
Db 373 TTTTAAAAAAACCTGTTTTTGCAAAAACGGTCTCAATCTTCATCTGTTCTTTTGTCT 432
QY 567 TACTCTCTTGGCTTAACATGATTTCTGACCAAGCAGCCGAGACAGACATGTGAGA 626
Db 433 TCATCTCCCTCCCAATATGATCTTAGCAACAAGAGACACATCTCTGTGAAAA 492
QY 627 AATGCTCTTCTTAATATGAGATTCGTAGTCTGGCATGAAATAGTAAATTAACATCT 686
Db 493 AGTGTCTCTTCTTAAGAGGGCTCTGTGGGCTGAATGCAATGATTAATTAACATAT 552
QY 687 GTCAAGTCAATTTCTGATTAATTTCT 713
Db 553 GCCAGTTTATTTCTGACTGTTTAT 579

RESULT 14

US-10-025-335-2

; Sequence 2, Application US/10025335
; Patent No. US20020127648A1

; GENERAL INFORMATION:

; APPLICANT: Coleman, Roger

; Bandman, Olga

; TITLE OF INVENTION: NOVEL HUMAN CSA-LIKE RECEPTOR

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/025,335

; FILING DATE: 18-Dec-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/791,974

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Billings, Lucy J.

; REGISTRATION NUMBER: 36,749

; REFERENCE/DOCKET NUMBER: PF-0198 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-855-0555

; TELEFAX: 415-845-4166

; TELEX: <Unknown>

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1257 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; IMMEDIATE SOURCE:

; LIBRARY: Consensus

; CLONE: 346874

; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-10-025-335-2

Query Match 11.1%; Score 131.8; DB 12; Length 1257;
Best Local Similarity 48.5%; Pred. No. 1.3e-24;
Matches 429; Conservative 1; Mismatches 443; Indels 12; Gaps 2;

QY 142 GAACACCACTCTGTGCACAGACATACAAATACCAACCCAGGTCCTTCCCACTGCTCTA 201

Db 208 GACAAACAGTTCCTTCTTCGCCAGTTTATTAAGATCTGGACCATTCATGATTTT 267
QY 202 CACTGTCTGTCTTTTGTGACTATACAAATGGCCGGCATGAGATTTCTTTCA 261
Db 268 TTTATTAGTTTCTCTGTGGAATTTATGGAATGTTTGTGCAACCTGGCTTTTATCA 327
QY 262 AATCCGAGTAATCAAACT--TTATTATTTTCTTAAGAACACAGTCAATTTGATCT 318
Db 328 GAAGATACGAATACACAGGTTGTGAGATCTACTTAATTAATTTGCTTACACGATTT 387
QY 319 TCTCATGATTCGTACTTTTCCATTCAAAATTTCTTATGATGCCAACTGGGAACAGACC 378
Db 388 CCTGCTTACTCTGGCAATTAACCAAGTAATAATGTTGTTGATTTGGGTGGCACCTTGGA 447
QY 379 ACTGAAACATTTTGTGTGACATTTACCTCCGATATTTTATTTTACAGATTAATTCAG 438
Db 448 RCTGAAGATATTCACACTGCCAGTAACAGCTCCCTCATCTATATCAATATGATTTATC 507
QY 439 TATTTCATTCCTGGGAGTATTAATCATATGATGCTACAGAACACCAAGCCATTTAA 498
Db 508 AATTATCTTCTTATGATTTGTGACATTTGACGCTGTCTTACAGCTGACACAGCTGAA 567
QY 499 AATATCCAAACCCAAATCTCTGGGGCTAAGATTTCTCTGTTGATCTGGGCATT 558
Db 568 GATCTACCGAATACAGAACCCGCTTCCCAAAATGATATCAACCGTTGTGTGCTAAT 627
QY 559 CATGTTCTTACTCTCTTGTGCTTACATGATTTGACCAACAGCAGCAGGAGACAGAA 618
Db 628 GGTCTCTTCTTATATGCTGCAAAATATGATGATTCCTCATCAAGATACAGAAAGTC 687
QY 619 TGTGAAGAAATGCTCTTCTTAATACAGATTCGCTAGTCTGAGTGAATAATAGTAA 678
Db 688 AATGTGGTGTGATGAGATTTAAAGAAATTTGAAAGAAATTTGCTGTGACAAA 747
QY 679 TTACATCTGTCAAGTATTTTCGGAATTAATTTCTTAATTTGATTTATGATATAC 738
Db 748 TTTCAATATGCTGACAAATATTTTAAATTTCTGAGCAATCAATTTAATTTCAATTTG 807
QY 739 CATTAACAAAGAACTGACGGTCAATGATGTAAGAACAGAGGGGTGTGTAAGTCCCA 798
Db 808 TGTAAATTCAGACGCTCTACAG-----AAACAAAGATTAATGAATAATTCACCAAT 858
QY 799 GAAAAAGGTGAACGCTCAAGATTTTCAATTAATGCTGATTTCTTATTTGTTTGTTC 858
Db 859 GAAAAAGCTCTCATCAATCAATTTTAAATTTCTGAGCAATCAATTTAATTTCAATTTG 918
QY 859 TTTCCATTTTGGCCGAATTCCTTACACCTGAGCCAAACCCGGAGTCTTGACGAC 918
Db 919 TTTACCATTTGTCGAAATCCGATACCTTCAGCCAGACAGAAATCAATGATGATGCTC 978
QY 919 TGTGAAATTAATCTGTTCTATGTGAAGAGACACTGTGTTTAACTTCTTAATATC 978
Db 979 AACCAAGATTTCACTTTCAAAAGCCAAAGAGGCTACAGCTGCTGCTGCTGCAACT 1038
QY 979 ATGCTGCAATCGGTATATTTTCTTCTTGAAGTCTTCAG 1023
Db 1039 GGTCTTGAATCTATCTTACTATCACTCAAAAGATTCGC 1083

RESULT 15

US-09-848-889-10

; Sequence 10, Application US/0984889
; Patent No. US2002002555A1

; GENERAL INFORMATION:

; APPLICANT: Au-Young, Janice

; Applicant: Guegler, Karl J.

; TITLE OF INVENTION: GPCR DIAGNOSTIC FOR BRAIN CANCER

; FILE REFERENCE: PC-0042 CIP

; CURRENT APPLICATION NUMBER: US/09/848,889

; CURRENT FILING DATE: 2001-05-03

; NUMBER OF SEQ ID NOS: 12

```
; SOFTWARE: PERL Program
; SEQ ID NO 10
; LENGTH: 518
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US2002002555A1 7979275H1
US-09-848-889-10
```

```
Query Match      10.7%; Score 126.6; DB 10; Length 518;
Best Local Similarity 57.8%; Pred. No. 1.9e-23;
Matches 225; Conservative 0; Mismatches 164; Indels 0; Gaps 0;
```

```
QY 152 CTGTGCACGAGAGACTACAAATACACCAGTCTCTCCACATGCTCTACACATGCTCTG 211
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 128 CGGTGCCCGACAGACCTCGAATGATGACAGCTGGTATTCGAGCCCTCTACACAGTGTT 187
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 212 TTTTGTGTGACTTATCACAATGGCCCTGGCGATGAGATTCTTTCAAATCCGAGT 271
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 188 TTCTTGACCGGCATCCCTGCTGAATACTTGTGCTGTGGGTGTTGTTACATCCCGCAGC 247
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 272 AAATCAACCTTATTTATTTCTTAAGACAGATCTATTTCTGATCTCTCATGATCTG 331
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 248 TCCTCCACCTTCACTACCTCAAAACACTTGTGGCGGACTTGATATATGACACTC 307
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 332 ACTTTCCATTCAAAATTTCTTAGTATGCCAACTGGGAACAGAGCCACTGAGAACTTTT 391
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 308 ATGCTTCTTTCAAAATCTCTCTGACTCACACCTGGCACCTGGCAGCTCAGAGCTTTT 367
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 392 GTGTGTCAGTTACCTCCGTCAATATTTTATTCACAATGTATATCAGTATTTCAATTCCTG 451
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 368 GTGTGTCTTTTCTCTGCGTATTTATGAGACATGTATGTGGGCATCGTGTGTGA 427
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 452 GGACTGTACTATTCGATCGCTACCAAGACAGCAGGCCATTTAAACATCCACACCC 511
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 428 GGGCTCATGACCTTTGACAGATTCTCCAGATCTCAGACCTTTGAGAAATATTTTCTA 487
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 512 AAAAATCTTGTGGGGCTAAGATTCTCTC 540
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 488 AAAAACCCTGTTTTTGCAAAAACGGTCTC 516
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: December 11, 2002, 09:13:09
Job time : 120 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 11, 2002, 03:09:10 ; Search time 2197 Seconds

(without alignments)
8750.133 Million cell updates/sec

Title: US-09-835-922-1

Perfect score: 1187
Sequence: 1 ggtaccggggagatctctag.....tgtgtcagaactcgtaaa 1187

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estda:*
2: em_estnum:*
3: em_estlin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vtl:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	732.2	61.7	1541	AK013804	AK013804 Mus muscu
2	732.2	61.7	2103	AK014807	AK014807 Mus muscu
3	638.2	53.8	937	BI758761	BI758761 603023751
4	623.8	52.6	680	BI758761	BI758761 602672256
5	616.8	52.0	1043	BI754688	BI754688 603025385
6	600	50.5	738	BI716135	BI716135 602677640

C	7	599.6	50.5	609	17	AO581896
	8	577.2	48.6	971	14	BO893760
	9	507.8	42.8	511	14	BM987216
	10	505.4	42.6	706	10	AV718096
	11	449.4	37.9	674	10	BB624828
	12	442.4	37.3	1681	14	BM926255
	13	424.2	35.7	877	14	BM919208
	14	407.6	34.3	542	9	AA274112
	15	366.8	30.9	915	12	BF677440
	16	334.8	28.2	476	10	AM045980
	17	307.2	25.9	1312	11	AF345565
	18	249	21.0	265	17	AO581861
	19	246.2	20.7	2515	11	AK008013
	20	240.8	20.3	257	14	BM713326
	21	240.8	20.3	262	14	BM674673
	22	212.6	17.9	775	11	BC029363
	23	208.2	17.5	1199	11	AF345566
	24	203	17.1	729	13	BI489709
	25	190.8	16.1	800	12	BM621718
	26	180.2	15.2	961	13	BI103980
	27	170.8	14.4	668	9	AU168581
	28	162.8	13.7	851	12	BM677820
	29	154.2	13.0	455	9	AI090920
	30	153.4	12.9	211	10	BE687381
	31	151.6	12.8	615	10	BB632376
	32	150.8	12.7	669	10	BB653415
	33	149.6	12.6	622	14	BM935684
	34	142.2	12.0	467	9	AT743546
	35	132	11.1	421	9	AI028704
	36	128.8	10.9	554	17	AO241005
	37	126.6	10.7	790	12	BG205036
	38	123.8	10.4	418	10	AW295791
	39	123.6	10.4	593	12	BE879604
	40	123	10.4	478	12	BF565992
	41	121.4	10.2	976	14	BO900494
	42	119.8	10.1	813	10	BE563611
	43	114.2	9.6	586	17	BM327804
	44	113.8	9.6	360	9	AI656746
	45	113.6	9.6	519	9	AI050350

ALIGNMENTS

RESULT 1
AK013804
LOCUS
DEFINITION Mus musculus adult male hippocampus cDNA, RIKEN full-length
RECEPTOR, full insert sequence.
ACCESSION AK013804
VERSION AK013804.1 GI:12851306
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclerogasthi; Muridae; Murinae; Mus.
1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
JOURNAL 99279253
MEDLINE 10349636
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
JOURNAL 20499374
MEDLINE
PUBMED 11042159

REFERENCE
AUTHORS

3 Shibata, K., Itoh, M., Aizawa, K., Nagata, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishii, K., Kitzunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kasliwal, K., Fujisake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yonekura, H., Ishikawa, T., Ozawa, K., Tanaka, T., Matsui, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer *Genome Res.* 10 (11), 1757-1771 (2000)

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Functional annotation of a full-length mouse cDNA collection
 Nature 409 (6821), 685-690 (2001)

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

5 (bases 10 to 1541)

Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bulic, C.,
Carrincci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T.,
Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hizaoka, T., Hori, F.,
Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T.,
Kato, H., Kawai, Y., Kojima, Y., Kono, H., Konda, M., Koya, S.,
Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K.,
Numata, R., Ono, M., Okeazaki, Y., Okido, T., Owa, C., Quackenbush, J.,
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
Schmickl, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,
Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamahaka, I.,
Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and
Hayashizaki, Y.

TITLE
JOURNAL

Submitted (0-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome@gsc.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222).

COMMENT

please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGGAGAGAGAGAGATCCAGAGCTTTTTTTTTTTTTTTNN 3'], cDNA was prepared by using thermostable thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGGAGAGAGATTCGATTATTTAAATTAATCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites 5' end: XhoI; 3' end: SstI.

Host: SOLR

FEATURES	Location/Qualifiers
source	1. .1541

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="FANTOM.DB:2900079B22"
/db_xref="MGD:MGI:1905984"
/db_xref="taxon:10090"
/clone="2900079B22"
/sex="male"
/tissue_type="hippocampus"
/clone_lib="RIKEN_full-length enriched mouse cdna library"
/dev_stage="adult"
358. .1401

```

CDS

```

/note="data source:SPR, source key:Q9DEPX4, evidence:ISS
homolog to P2Y12 PLATELET ADP RECEPTOR
putative"
/codon_start=1
/protein_id="BAB29000.1"
/db_xref="GI:12851307"
/db_xref="MGD:MGI:1920308"
/translation="ADYGVNTTSANTTFSPGTSILCYRDYKKTIOVLEPPLIYLYLFF
GLITSLAMRIFFQILRSKSNFIFLEKNTVISDLMILFEPKLLISDAVLGAPRLTIV
GVNVSVEFYEMYSISFLGITLIDRYLKTTPFTSSPMLGAKLISVYTAFMFVPL
ISLPMILNRPKDQVTKCSFLSEEGALWHELVNVCQVIFININLIYVICVSLI
TKELRSVTRPGRGSAKPKKKNVYKVFILIAVFICPPFPHARIPYLSOTRACED
SAENTLFVKESTLMLTSLNACLDPIFFLEFKCSRNLSLMLRCSNSTSGNKKK

```

polyA_signal

```
polyA_site      /note="putative"
1541
```

BASE COUNT	431 a	384 c	296 g	430 t
ORIGIN				

Query Match	61.7%;	Score 732.2;	DB 11;	Length 1541;
Best Local Similarity	82.2%;	Pred. No. 5e-149;		
Matches 866; Conservative	0;	Mismatches 183;	Indels 4;	Gaps 2;

114 CCGTGCACATCTCACCTCTGCGCTGGGAACACACAGTCTGTGCACACAGACTACAAA 173

383 CCTCAGCAATACCAACCTTCTCCCCGCGGACCAAGACCCCTGTGCGTCAGAGACTTACAAGA 442

174 TCACCCAGTCTCTTCCCACTGCTACACATGTCCTGTTTTTTGTTGGACTTATCACAA 233

443 TCACCCAGGTTCTTCCCATTTGCTGTACACCCGTCCTGTCTTTGCTGGGCTCATCAGCA 502

234 ATGGCCTGGCGATGAGGATTTCTTTCAATCCGGAGTAATCAAACTTTATTATTTTTC 293

503 ACAGCTGGCAATGAGGATTTTCTTTTCAGATCCGAGTAATAATCCAACTTCATCATTTTTTTTTC 562

294 TTTAAGAACACAGTCATTTCGTATCTCTCATGATTCGACCTTTTCATTCAAAAATTCCTTA 353

[illegible]

354 GTTGAATGCCAAAACCTGGGGGACACAGGACCACTGACAGAACTGTTTGTGCTCTCTCAACACCTTAACTCTGCTCTCA A13

[illegible][illegible]

414 TATTTATTTTCCAAATGATATATCAGTATTTTCAATCCGGAGCTGATTAACATATGATTCCT 4/3

683 CATTCTATTACCAAGTATATACAGTATATCGTTCCTGGGCTGTGATACCATTTGACCCGT 742

4/74 ACCGAGAGCCACCAGGCCATTTAAACATCCACCCCCAAAATCTCTGGGGGCTAGA 533

743 ACCTGAGACCCAGGCCATTAAACGTCACAGCCCGAGCATCTCTGGGTGCAAGA 802

534 TTCTCTCTGTTGTCATCTGGGCATTCATGTTCCTTACTCTCTTTGCCATAACATGATTTCTGA 593

803 TTCTTCTGTGTCATCTGGCCCTTCATGTTCTTAATTCTACTGCTTACATGATTCTCA 862

594 CCACAGCGACGCCGAGAGACCAAGAAATGTGAAGAAATGCTCTTTCTTTAAATCAGAGTTTCG 653

Db	863	CCAACAGGAGGCCAAAAGATTAAGAGCTAACAAAAGTCTTTCTTAAGTCAGAGTTG	922
OY	654	GTCAGTCTGGCATGAATAGTAAATTACATCTGTCAAGTCATTTTCTGGATTAATTCT	713
Db	923	GTCATGATTTGGCAGAAATAGTCATTAATTCATCTGCCAAGTCATTTCTGGATTATTTT	982
OY	714	TAAATGTTATTTGATGTTATACACTATTAACAAAAGTACAGCGGTACATCGTAAGA	773
Db	983	TAAATGTCATCGTTTGTTATTAAGCTCATTAACAAAACATCATGCTTATGTACAGAA	1042
OY	774	CGAGGGGTGAGGTAAGATCCCGAGAAAAAGGTGAACATCAAGTTTTCATTATCATTG	833
Db	1043	CNAGGGTTTACGCCAAGATCCCAAGAAAAAGGTAAAGTCAAGTTTTCATCATCATTG	1102
OY	834	CTGATTTCTTATTTTGTGTGTTCCATTTCATTTTGGCCGAATTCCTTACACCCGAGCC	893
Db	1103	CTGATTTCTTATTTTGTGTTGTTCCATTTCATTTTGGCCGAATTCCTTACACCCGAGCC	1162
OY	894	AAACCCGGGATGTCCTTTGACTGCATCTGCTGAAATATCTGTGTTCTATGTGAAGAGCA	953
Db	1163	AAATCTCGGGCGCTTTTGACTGTCGAGCTGAGAACACCTGTGTTCTACGATGAAGGAGACA	1222
OY	954	CTCTGCTGTTAACTTCCTTAAATGCAATGCTGGATCCGTCATTCATTTTTCCTTGCA	1013
Db	1223	CCCATGAGCTGACGTACTGTAAGCCCTGCTGTATCATTCATCTACATTTTCTTTCGA	1282
OY	1014	AGTCTCTCAGAAAAATCCCTGATTAAGTATGCTGAAGTGCACCAATTCGCAACATCTGT	1073
Db	1283	AGTCTTTCAGAAATTCCTTATACAAAGCATGCTGAGTGTCTCAAACTCT---ACATCAACAT	1353
OY	1074	CCCAAGCAATAGGAAAAAAGAACAGATGCTGTGACCCCAATGAGAGACTCCATGT	1133
Db	1340	CTGGGCAAAACAAAGAAAGAGACAAAGAGTGGCAACCAAGCAAGAGACCCCAATGT	1399
OY	1134	AAACAATTAACATAGGAATATTTCAATCTCT	1166
Db	1400	AGA-ACATTAACCAAGGGGCTGCTGTACGCTTT	1431

	RESULT 2
	AK014807
LOCUS	
DEFINITION	AK014807 2103 bp mRNA linear HTC 19-JAN-2002
ACCESSION	Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4921504D23;putativegic receptor PZy, G-protein coupled 12, full insert sequence.
VERSION	AK014807
KEYWORDS	AK014807.1 GI:12852867
SOURCE	HTC; CAP trapper. Mus musculus (Strain:C57BL/6J) adult male testis cDNA to mRNA, clone_1lib:RIKEN full-length enriched mouse cDNA library clone:4921504D23.
ORGANISM	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1 Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999)
AUTHORS	JOURNAL MEDLINE 99279253 PUBMED 10349636
REFERENCE	2 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)
JOURNAL MEDLINE	20499374 PUBMED 11042159
REFERENCE	3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Kono,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Suni,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegai,T., Kasliwalgi,K.,

TITLE	JOURNAL	PUBMED	REFERENCE
Fujisake, S., Inoue, K., Togawa, Y., Izawa, M., Ohura, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, K., Kira, A. and Hayashizaki, Y.	RIKEN Integrated sequence analysis (RISA) system -384-Format		
sequencing pipeline with 384 multicapillary sequencer	Genome Res.	10 (11), 1757-1771	(2000)
Kawai, J., Shitagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,		4	
11076861			

TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409 (6821), 685-690 (2001)
MEDLINE	21085660
PUBMED	11217851
REFERENCE	5 (bases 1 to 2103)
AUTHORS	Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,

TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suenhiro cho, Tsukumi-ku, Yokohama Kanagawa 230-0045, Japan (E-mail: genome-res@gs.c.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT Please visit our web site (<http://genome.gsc.riken.go.jp/>) for

FEATURES	Location/Qualifiers
SOURCE	1. .2103

/db_xref="taxon:9606"
 /clone_lib="NIH_MGC_114"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMV-Sport6; Site: 1; NotI; site_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27 yo. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 019. Note: this is a NIH_MGC Library."
 BASE COUNT 273 a 227 c 164 g 272 t 1 others
 ORIGIN

Query Match 53.8%; Score 638.2; DB 13; Length 937;
 Best Local Similarity 94.0%; Pred. No. 1.4e-128;
 Matches 719; Conservative 0; Mismatches 38; Indels 8; Gaps 5;

QY 28 CTGAGGCTGCAATTAATCTACTTACTGATACATTCAAACCTCCAGATCAACAGT 87
 Db 174 CTGAGGCTGCAATTAATCTACTTACTGATACATTCAAACCTCCAGATCAACAGT 233
 QY 88 ATCAGGTACCAACAAATGCAAGCCGTCGACATCTCCTGCGCTGGGAACAC 147
 Db 234 ATCAGGTACCAACAAATGCAAGCCGTCGACATCTCCTGCGCTGGGAACAC 293
 QY 148 CAGTCTGTGACACGAGACTACAAATCACCAGGTCTCTCCAGTCTACAGCT 207
 Db 294 CAGTCTGTGACACGAGACTACAAATCACCAGGTCTCTCCAGTCTACAGCT 353
 QY 208 CCGTGTGTGTGGAGCTTATACAAATGGCTGGCATGAGATTTCTTTCAATCCG 267
 Db 354 CCGTGTGTGTGGAGCTTATACAAATGGCTGGCATGAGATTTCTTTCAATCCG 413
 QY 268 GAGTAATCAAACTTTATTTTCTTAAACACAGTCTATTTCTGATCTTCATGAT 327
 Db 414 GAGTAATCAAACTTTATTTTCTTAAACACAGTCTATTTCTGATCTTCATGAT 473
 QY 328 TCTGACTTTTCCATTCAAATTTCTTATGATGCCAAACTGGGAACAGACACTGAGAC 387
 Db 474 TCTGACTTTTCCATTCAAATTTCTTATGATGCCAAACTGGGAACAGACACTGAGAC 533
 QY 388 TTTTGTGTGCAAGTACCTCCGTCATATTTTATTTTCAAAATGATATACATTTTCAT 447
 Db 534 -TTTGTGTGCAAGTACCTCCGTCATATTTTATTTTCAAAATGATATTTTCAT 592
 QY 448 CTTGGGACTGATACATTCATGCTGACGAGAACACACAGGCCATTTAAACATCCAA 507
 Db 593 CTTGGGACTGATACATTCATGCTGACGAGAACACACAGGCCATTTAAACATCCAA 652
 QY 508 CCCCCAAATCTCTGGGGGCTAAGATTTCTGTGTGCACTGGGATTCATGTTCTT 567
 Db 653 CCCCCAAATCTCTGGGGGCTAAGATTTCTGTGTGCACTGGGATTCATGTTCTT 712
 QY 568 ACTCTCTTGGCTAACATGATTTGACCAACAGCAGCCGAGAGACAAGATTTGAAG-A 626
 Db 713 ACTCTCTTGGCTAACATGATTTGACCAACAGCAGCCGAGAGACAAGATTTGAAGAA 772
 QY 627 AATTCCTTTTCTTAAATACAGAGTTCCG--TCTAGTCTGGGATGAATATGATATACAT 684
 Db 773 AATTCCTTTTCTTAAATACAGAGTTCCG--TCTAGTCTGGGATGAATATGATATACAT 832
 QY 685 CTGTCAGTCAATTTCTGATTAATTTCTTAAATGATTTATGATATACATCAATAC 744
 Db 833 CTGTCAGTCAATTTCTGATTAATTTCTTAAATGATTTATGATATGATTTGATTA--CCCTCCTTAC 889
 QY 745 AAAAGAACTGATACGATACAGTAAGAACGAGGGGTGTAGTAA 789
 Db 890 CAAAGAACTGGA-CGGTCATACGTAAGAACCAAGGGGTTTGTATAA 933

RESULT 4
 BG708321 680 bp mRNA linear EST 07-MAY-2001
 LOCUS 602672256F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4795236 5',
 DEFINITION mRNA sequence.
 ACCESSION BG708321
 VERSION BG708321.1 GI:13985546
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 680)
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: g9abs-remail.nih.gov
 Tissue Procurement: Miklos Palakivits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shihaki
 Toshiyuki and Piero Carninci (RIKEN)
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 found through the I.M.A.G.E. Consortium information can be
 http://image.llnl.gov
 Plate: LHAM10677 row: 0 column: 13
 High quality sequence stop: 679.
 location/Qualifiers
 1. 680
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="IMAGE:4795236"
 /clone_lib="NIH_MGC_96"
 /tissue_type="hypothalamus"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pBluescript (modified pBluescript KS+); Site: 1; BamHI; Site: 2; SalI-XhoI (glcagag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTTTTT-3', size-selected for average insert size 2.3 kb and normalized to R0.5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."
 BASE COUNT 198 a 169 c 118 g 195 t
 ORIGIN

Query Match 52.6%; Score 623.8; DB 12; Length 680;
 Best Local Similarity 99.5%; Pred. No. 2e-125;
 Matches 636; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 28 CTGAGGCTGCAATTAATCTACTTACTGATACATTCAAACCTCCAGATCAACAGT 87
 Db 33 CTGAGGCTGCAATTAATCTACTTACTGATACATTCAAACCTCCAGATCAACAGT 92
 QY 88 ATCAGGTACCAACAAATGCAAGCCGTCGACATCTCCTGCGCTGGGAACAC 147
 Db 93 ATCAGGTACCAACAAATGCAAGCCGTCGACATCTCCTGCGCTGGGAACAC 152
 QY 148 CAGTCTGTGACACGAGACTACAAATCACCAGGTCTCTCCAGTCTACAGCT 207
 Db 153 CAGTCTGTGACACGAGACTACAAATCACCAGGTCTCTCCAGTCTACAGCT 212
 QY 208 CCGTGTGTGTGGAGCTTATACAAATGGCTGGCATGAGATTTCTTTCAATCCG 267
 Db 213 CCGTGTGTGTGGAGCTTATACAAATGGCTGGCATGAGATTTCTTTCAATCCG 272
 QY 268 GAGTAATCAAACTTTATTTTCTTAAACACAGTCTATTTCTGATCTTCATGAT 327
 Db 273 GAGTAATCAAACTTTATTTTCTTAAACACAGTCTATTTCTGATCTTCATGAT 332
 QY 328 TCTGACTTTTCCATTCAAATTTCTTATGATGCCAAACTGGGAACAGACACTGAGAAC 387

```

|||||
Db 333 TGTGATTTTTCATTCATAAAATTCCTTAGTATGCGCAAACTGGGAACGAGACCACTGAGAAC 392
Oy 388 TTTTGTGTCAAGTACCTCCGTATTTTATTTTACAAATGTATATGATTTTCATT 447
Db 393 TTTTGTGTCAAGTACCTCCGTATTTTATTTTACAAATGTATATGATTTTCATT 452
Oy 448 CCTGGACGTATATGATGATGCTACGAGAAAGACCAGGCCATTTAAACATCCAA 507
Db 453 CTGGGACGTATATGATGATGCTACGAGAAAGACCAGGCCATTTAAACATCCAA 512
Oy 508 CCCCCAAATCTCTGGGGGGCTAAGATCTCTCTGTGTGCATCTGGGCATTCATCTCTT 567
Db 513 CCCCCAAATCTCTGGGGGGCTAAGATCTCTCTGTGTGCATCTGGGCATTCATCTCTT 572
Oy 568 ACTCTCTTGCCTAACATATTTCTGACCAAGCAGCAGCAGAGCAAAATGTGAAG-CA 626
Db 573 ACTCTCTTGCCTAACATATTTCTGACCAAGCAGCAGCAGAGCAAAATGTGAAGCA 632
Oy 627 AATGCTCTTCTTAATCAGAGTTCGGTCTAGTCTGGC 665
Db 633 AATGCTCTTCTTAATCAGAGTTCGGTCTAGTCTGGC 671

```

```

RESULT 5
LOCUS B1754688 1043 bp mRNA linear EST 25-SEP-2001
DEFINITION 603025383F1 NIH_MGC_114 Homo sapiens CDNA clone IMAGE:5195900 5',
ACCESSION B1754688
VERSION B1754688
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

```

```

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-riemail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed By: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM11490 row: e column: 21
High quality sequence stop: 733.
Location/Qualifiers
1..1043
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="5195900"
/clone_id="NIH_MGC_114"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site: 1; NotI;
Site: 2; EcoRV (destroyed); RNA source anonymous pool of 6
male brains; age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 019. Note:
this is a NIH_MGC Library."

```

```

BASE COUNT 340 a 228 c 205 g 265 t 1 others
ORIGIN

```

```

Query Match 52.0%; Score 616.8; DB 13; Length 1043;
Best Local Similarity 99.4%; Pred. No. 6-2e-124;
Matches 640; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

```

```

Oy 544 TGTATCTGGGCAATTCATCTGTCTTACTCTCTTGGCCTAACATGATTTGACCAAGGCA 603
Db 13 TGTATCTGGGCAATTCATG-CTTACTCTC-TTGCTAACATGATTTCTGACCAAGGCA 70
Oy 604 GCCGAGACCAAGATGTGAAGAAATGCTCTTCTTAAATCAGAGTTGGTCTACTCG 663
Db 71 GCCGAGACCAAGAAATGTGAAGAAATGCTCTTCTTAAATCAGAGTTGGTCTACTCG 130
Oy 664 GCATGAATAGTAAATTAATCATCTGTCAAGTCAATTTCTGAGTAAATTTCTAATGTTAT 723
Db 131 GCATGAATAGTAAATTAATCATCTGTCAAGTCAATTTCTGAGTAAATTTCTAATGTTAT 190
Oy 724 TGTATGTTATACACTTATTAACAAGAACTGTAACCGGTATACGTAAGAAGAGGGGTCT 783
Db 191 TGTATGTTATACACTTATTAACAAGAACTGTAACCGGTATACGTAAGAAGAGGGGTCT 250
Oy 784 AGTAAAGTCCCGAGAGAAAGGTGAAGCTCAAAAGTTTCATTTATTCATTTGCTGATCTT 843
Db 251 AGTAAAGTCCCGAGAGAAAGGTGAAGCTCAAAAGTTTCATTTATTCATTTGCTGATCTT 310
Oy 844 TATTTGTTTGTCTCTTCCATTTTGGCGAATTCCTTACCCCTGAGCCAAACCGGGA 903
Db 311 TATTTGTTTGTCTCTTCCATTTTGGCGAATTCCTTACCCCTGAGCCAAACCGGGA 370
Oy 904 TGTCTTTGACTGCACCTGTAAGAAATACCTCTGTTCTATGTGAAGAGAGACACTGTGCTT 963
Db 371 TGTCTTTGACTGCACCTGTAAGAAATACCTCTGTTCTATGTGAAGAGAGACACTGTGCTT 430
Oy 964 AACTTCCTTAATGATGATGCTGCTGATCCGTTCATCTATTTTCTTCTTTCGAACTCTTCAG 1023
Db 431 AACTTCCTTAATGATGATGCTGCTGATCCGTTCATCTATTTTCTTTCGAACTCTTCAG 490
Oy 1024 AATTCCTGATATGATGATGCTGATGCTGCTGATGCTGCTGATGCTGCTGCTGCTGCT 1083
Db 491 AATTCCTGATATGATGATGCTGATGCTGCTGATGCTGCTGATGCTGCTGCTGCTGCT 550
Oy 1084 TAGGAAAAAAGAACAGAGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 1143
Db 551 TAGGAAAAAAGAACAGAGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 610
Oy 1144 ACTAAGGAATATTTCAATCTCTTGTGTCAGAACTGCTTAA 1187
Db 611 ACTAAGGAATATTTCAATCTCTTGTGTCAGAACTGCTTAA 654

```

```

RESULT 6
LOCUS BG716155 738 bp mRNA linear EST 08-MAY-2001
DEFINITION 602677640F1 NIH_MGC_96 Homo sapiens CDNA clone IMAGE:4800249 5',
ACCESSION BG716155
VERSION BG716155
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-riemail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed By: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM10690 row: p column: 10

```


High quality sequence start: 2
High quality sequence stop: 706.

FEATURES
Location/Qualifiers

1..738

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="IMAGE:480249"

/clone_lib="NIR_MGC_96"

/tissue_type="hypothalamus"

/lab_host="DH10B"

/note="Organ: brain; Vector: pBluescript (modified pBluescript KS+); Site_1: BamHI; Site_2: SalI; XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTT-3', size selected for average insert size 2.3 kb and normalized to 10⁶ 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carlnaci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIR_MGC Library."

BASE COUNT 210 a 184 c 126 g 217 t 1 others

ORIGIN

Query Match 50.5%; Score 600; DB 12; Length 738;

Best Local Similarity 95.3%; Pred. No. 2.9e-120;

Matches 672; Conservative 0; Mismatches 26; Indels 7; Gaps 5;

```

QY 28 CTCGAGGCTGCAATTAACCTACTACTGATGATCAATCAACCCCTCCAGATCAACAGTT 87
DB 34 CTCGAGGCTGCAATTAACCTACTACTGATGATCAATCAACCCCTCCAGATCAACAGTT 93
QY 88 ATAGAGTAACCAACAGAAATGCAAGCCGTCGACATCTCCTCGCCCTGGGAACAC 147
DB 94 ATAGAGTAACCAACAGAAATGCAAGCCGTCGACATCTCCTCGCCCTGGGAACAC 153
QY 148 CAGTCTGTCACGACGACATCAAAATCACCAGGCTCTCCACTGCTACACTGT 207
DB 154 CAGTCTGTCACGACGACATCAAAATCACCAGGCTCTCTCCACTGCTACACTGT 213
QY 208 CCGTCTTTTGTGGAATTCACAAATGGCTGGCATGAGATTTCTTCAATCCG 267
DB 214 CCGTCTTTTGTGGAATTCACAAATGGCTGGCATGAGATTTCTTCAATCCG 273
QY 268 GAGTAATCAAACTTATTTTCTTAAGAACACAGTATTTGATCTTCATGAT 327
DB 274 GAGTAATCAAACTTATTTTCTTAAGAACACAGTATTTGATCTTCATGAT 333
QY 328 TCGACTTTCATCAATCAATTTCTAGTATGCAAACTGGGAACAGACAGACGAGAAC 387
DB 334 TCGACTTTCATCAATCAATTTCTAGTATGCAAACTGGGAACAGACAGACGAGAAC 393
QY 388 TTTTGTGTCAAGTACCTCCGTCATATTTTATTCACATGTA--TATCAGTATTCAT 446
DB 394 TTTTGTGTCAAGTACCTCCGTCATATTTTATTCACATGTAATGATTTTCAT 453
QY 447 TCGTGGAGCTGATACATGATGCTACCAAGAACACAGGCCATTTAAACATCCA 506
DB 454 TCGTGGAGCTGATACATGATGCTACCAAGAACACAGGCCATTTAAACATCCA 513
QY 507 ACCCCAA- AATGCTCTGGGGGCTAAGATTCCTCTGTGTCAATCTGGGATTCATGTC 565
DB 514 NACCCAAAGAAATCTCTGGGGGCTAAGATTCCTCTGTGTCAATCTGGGATTCATGTC 573
QY 566 TTAATCTCTTGGCTAATCATGATTCGACCAACAGGACGCGAGAGAACAA--TGTGA 623
DB 574 TTAATCTCTTGGCTAATCATGATTCGACCAACAGGACGCGAGAGAACAACTGTGAC 633
QY 624 AGAAATGCTCTTCTTAATCAAGAGTTGGG- TCTACTGTGGC--ATGAATAGTAATTT 680
DB 634 AGAAATGCTCTTCTTAATCAAGAGTTGGGATCTAGTCGGCATGAACATGTAATTT 693
QY 681 ACAATGCTCAAGTATTTCTGATTAATTTCTTAATTTGATTTG 725
DB 694 ACAATGCTCAAGTATTTCTGATTAATTTCTGATTTG 738

```

RESULT 7

LOCUS A0581896/c

DEFINITION RPCI-11-433N7.TV RPCI-11 Homo sapiens genomic clone RPCI-11-433N7,

DNA sequence.

ACCESSION A0581896

VERSION A0581896.1

KEYWORDS GI:5009006

SOURCE GSS.

ORGANISM human.

REFERENCE Homo sapiens

AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

1 (bases 1 to 609)

Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter

J.C.

TITLE use of BAC End Sequences from library RPCI-11 for Sequence-Ready

Map Building

Unpublished (1997)

Other-GSSs: RPCI-11-433N7.TV

Contact: Shaying Zhao, William Nierman, Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850

Tel: 301 838 0200

Fax: 301 838 0208

Email: hbe@tigr.org

Clones are derived from the human BAC library RPCI-11. For BAC

library availability, please contact Pieter de Jong

(pieterdejong.med.bufileo.edu). Clones may be purchased from

BACPAC Resources (<http://bacpac.med.bufileo.edu/ordering>) or from

Research Genet cs (inforesgen.com). BAC end search page:

http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html.

Seg primer: T7

Class: BAC ends.

FEATURES
Location/Qualifiers

1..609

/organism="Homo sapiens"

/db_xref="GDB:7665206"

/db_xref="taxon:9606"

/clone="RPCI-11-433N7"

/clone_lib="RPCI-11"

/sex="Male"

/cell_type="Lymphocytes"

/note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;

RPCI11 Human Male BAC Library"

BASE COUNT 204 a 104 c 122 g 179 t

ORIGIN

Query Match 50.5%; Score 599.6; DB 17; Length 609;

Best Local Similarity 99.3%; Pred. No. 3.6e-120;

Matches 602; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

QY 261 AATCCGAGTAATCAACTTATTTATTTCTTAAGAACACAGTCAATTCGATCTC 320
DB 609 AATCCGAGTAATCAACTTATTTATTTCTTAAGAACACAGTCAATTCGATCTC 320
QY 321 TCAATGATCGACTTTCATCAAAATTCCTATGATGCAAACTGGGAACAGGACAC 380
DB 549 TCAATGATCGACTTTCATCAAAATTCCTATGATGCAAACTGGGAACAGGACAC 380
QY 381 TGAGAACTTTGTGTCAAGTACCTCCGTCATATTTTATTCACATGTAATTCAGTA 440
DB 489 TGAGAACTTTGTGTCAAGTACCTCCGTCATATTTTATTCACATGTAATTCAGTA 440
QY 441 TTTTATCTCTGGGCTGATTAATGATGCTACCAAGAACACAGGCCATTTAA 500
DB 429 TTTTATCTCTGGGCTGATTAATGATGCTACCAAGAACACAGGCCATTTAA 500
QY 501 CATTCAACCCCAAAATCTCTGGGGCTAAGATTCCTCTGTGTCAATCTGGGCTATCA 560
DB 369 CATTCAACCCCAAAATCTCTGGGGCTAAGATTCCTCTGTGTCAATCTGGGCTATCA 310

```

QY 561 TGTTCTACTCTCTTTGGCTTAACATGATTCGACCAACAGGAGCCGAGACCAAGATG 620
 DB 309 TGTTCTACTCTCTTTGGCTTAACATGATTCGACCAACAGGAGCCGAGACCAAGATG 250
 QY 621 TGAAGAAAGCTCTTCCCTTAATCAGAGTGGCTCTACTCGGATGAAATAGTAAT 680
 DB 249 TGAAGAAAGCTCTTCCCTTAATCAGAGTGGCTCTACTCGGATGAAATAGTAAT 190
 QY 681 ACATCTGTCAGTCAATTTCTGATTAATTTCTTAATTTGATGATTTATACACTCA 740
 DB 189 ACATCTGTCAGTCAATTTCTGATTAATTTCTTAATTTGATGATTTATACACTCA 130
 QY 741 TTACAAAGAACTGTACCGGTACATACGTAAAGAGGGGTAGTAAAGTCCCGAGA 800
 DB 129 TTACAAAGAACTGTACCGGTACATACGTAAAGAGGGGTAGTAAAGTCCCGAGA 70
 QY 801 AAAGGTGAGGTCAAGTTTCATATCATGCTGATCTTATTTTGTGTTCTCT 860
 DB 69 AAAGGTGAGGTCAAGTTTCATATCATGCTGATCTTATTTTGTGTTCTCT 10
 QY 861 TCCATT 866
 DB 9 ACCATT 4
 RESULT 8
 LOCUS B0893760 971 bp mRNA linear EST 16-AUG-2002
 DEFINITION AGENCOURT.8482519 lupski_dorsal_root_ganglion Homo sapiens CDNA
 accession B0893760
 version B0893760.1 GI:22285774
 keywords EST.
 source human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 971)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: Dr. James R. Lupski
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L14M13577 row: m column: 06
 High quality sequence stop: 607.
 Location/Qualifiers
 1..971
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6186029"
 /clone_lib="lupski_dorsal_root_ganglion"
 /sex="male"
 /tissue_type="dorsal root ganglia"
 /dev_stage="adult, 36 yr"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPORT6 (Life Technologies); Site_1:
 NotI; Site_2: SalI; cDNA made by oligo-dT priming.
 directionally cloned using the following adaptors:
 5'-TCGACCCAGCGCTCCG-3' and
 5'-GACTAGTCTAGATCGAGCGCGCCCTT(15)-3'. Size selected >
 1 kb for average insert length 1.7 kb. This is a primary
 library, non-amplified. Library constructed by Life
 Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
 College of Medicine) and is available through Life
 Technologies."

BASE COUNT 299 a 236 c 172 g 264 t
 ORIGIN
 Query Match 48.6%; Score 577.2; DB 14; Length 971;
 Best Local Similarity 97.7%; Pred. No. 2.5e-115;
 Matches 596; Conservative 0; Mismatches 13; Indels 1; Gaps 1;
 QY 28 CTGCAAGCTGCAATTAAGTACTACTGATGATGATGATGATGATGATGATGAT 87
 DB 148 CTGCAAGCTGCAATTAAGTACTACTGATGATGATGATGATGATGATGATGAT 207
 QY 88 ATCAGTAAACCAAGAAATGCAAGCCGTGACAAATCTACCTCTGGGCGGAGAAC 147
 DB 208 ATCAGTAAACCAAGAAATGCAAGCCGTGACAAATCTACCTCTGGGCGGAGAAC 267
 QY 148 CAGTCTGTCACAGAGACTACAAATACCCAGGCTCTCCACAGTCTCTACAGT 207
 DB 268 CAGTCTGTCACAGAGACTACAAATACCCAGGCTCTCTCCACAGTCTCTACAGT 327
 QY 208 CCGTGTGTTTGTGAGCTTATGCAAAATGCGCTGGCGATGAGGATTTCTTCAATCG 267
 DB 328 CCGTGTGTTTGTGAGCTTATGCAAAATGCGCTGGCGATGAGGATTTCTTCAATCG 387
 QY 268 GAGTAATCAAACTTATTTATTTCTTAAAGACAGTCAATTTCTGATCTTCTATGAT 327
 DB 388 GAGTAATCAAACTTATTTATTTCTTAAAGACAGTCAATTTCTGATCTTCTATGAT 447
 QY 328 TCTGACTTTTCCATTCAAATTTCTAGTATGCGCAATGCGGAGACAGCAGTGAAG 387
 DB 448 TCTGACTTTTCCATTCAAATTTCTAGTATGCGCAATGCGGAGACAGCAGTGAAG 507
 QY 388 TTTGTGTGTCAAGTACCTCGCTCATATTTATTTCAACATGATATTCATTTCAAT 447
 DB 508 TTTGTGTGTCAAGTACCTCGCTCATATTTATTTCAACATGATATTCATTTCAAT 567
 QY 448 CCGTGGAGCTGATATATGATGATGATGATGATGATGATGATGATGATGATGAT 507
 DB 568 CCGTGGAGCTGATATATGATGATGATGATGATGATGATGATGATGATGATGAT 627
 QY 508 CCCCAAAATCTCTTGGGGGCTAAGATTCCTCTCTTGTGATCTGGGCTTCATGTTCT 567
 DB 628 CCCCAAAATCTCTTGGGGGCTAAGATTCCTCTCTTGTGATCTGGGCTTCATGTTCT 687
 QY 568 ACTCTCTTGTGCTTAACATGATTTCTGACACACA-GGCAGCCGAGAGCAAGATGGAACA 626
 DB 688 ACTCTCTTGTGCTTAACATGATTTCTGACACACA-GGCAGCCGAGAGCAAGATGGAAGA 747
 QY 627 AATGCTCTT 636
 DB 748 AATGCTCTT 757
 RESULT 9
 LOCUS BM987216 511 bp mRNA linear EST 25-MAR-2002
 DEFINITION UT-H-CO0-aql-c-12-0-01.s1 NCI_CGAP_Sud9 Homo sapiens CDNA clone
 accession BM987216
 version BM987216.1 GI:19706605.
 keywords EST.
 source human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 511)
 AUTHORS NCI-CGAP http://www.nci.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: Dr. Jose Mercuende
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNI at: <http://image.llnl.gov>
Seq primer: M13 FORWARD
SOLTA=Yes.

FEATURES	Location/Qualifiers
source	1. .511

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE: 3104254"
/clone_id="NCI_CGAP_Sub9"
/tissue_type="mixed"
/dev_stage="mixed"
/lab_host="DH10B (Life Technologies)"
/notes="Vector: pT73-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoR I; Site_2: Not I; tissues:
Cholonic mucosa with Crohns disease, Cholonic mucosa with
ulcerative colitis, Fetal thymus, Cervix, Cervical
adenosquamous carcinoma, Ligament cells, prostate
carcinoma, Bladder carcinoma, Brain oligodendroga
; NCI_CGAP_Sub9 is a subtraced CDNA library constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand CDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (GT)18 tail. The
sequence tags for this library are CCGC, AAGC, GGCC,
GGAG, TAGC, TTAGC, ATGG, AGAC, ATCAC. For additional
information, contact: Bento Soares, bento-soares@uiowa.edu
TAG_LIB=UI-H-COO
TAG_TISSUE=Brain Oligodendroga
TAG_SEQ=ATCAC"

```

Query Match	42.8%	Score 507.8	DB 14	Length 511
Best Local Similarity	99.6%	Pred. No. 3.3e-100		
Matches 509	Conservative 0	Mismatches 2	Indels 0	Gaps 0
QY 281	TTTATTATTTTCTTCTTACGACACAGCTCATTTCTGATCTCTCATGATATTTCTGACTTTTCCA	340		
Db 1	TTTTTTTTTTTTTCTTAAGAACACAGTCAATTTCTGATCTCTCATGATATTTCTGACTTTTCCA	60		
QY 341	TTCAAAATTTCTTAGTGATGCCAAACTGGGAGACGACCACTGAGAACTTTTGTTGTCA	400		
Db 61	TTCAAAATTTCTTAGGATGCCCAACTGGGAGACGACCACTGAGAACTTTTGTTGTCA	120		
QY 401	GTTACCTCCGTCATTTTTTTATTTCCAGTATTTATCAGATATTTCAATTCCTGGGACTGATA	460		
Db 121	GTTACCTCCGTCATTTTTTTATTTCCAGTATTTATCAGATATTTCAATTCCTGGGACTGATA	180		
QY 461	ACTATCGATCGCTACAGACAGACACCAAGGCCATTTTAAACATCCAACCCAAAAATCTC	520		
Db 181	ACTATCGATCGCTACAGACAGACACCAAGGCCATTTTAAACATCCAACCCAAAAATCTC	240		
QY 521	TTGGGGGCTAGATTTCTCTCTGTGTCTATCTGGCATTCATATGTTCTTACTCTCTTGGCT	580		
Db 241	TTGGGGGCTAGATTTCTCTCTGTGTCTATCTGGCATTCATATGTTCTTACTCTCTTGGCT	300		
QY 581	AACATGATTTGACCAACAGCGCAGCGGAGACAGAAATGGAAGAAATGCTTTCCTT	640		
Db 301	AACATGATTTGACCAACAGCGCAGCGGAGACAGAAATGGAAGAAATGCTTTCCTT	360		
QY 641	AAATCAGATTCGCTCTAGTCTGGCATGAAAATGATAATTTACATCTGTCAAGTCAATTTTC	700		
Db 361	AAATCAGATTCGCTCTAGTCTGGCATGAAAATGATAATTTACATCTGTCAAGTCAATTTTC	420		
QY 701	TGGATTAATTTCTTAATGTATTTGTATGTTATACACTCATTTACAAAAGACTGTACGG	760		

Db	421	TGATTAAATTTCTTAATATGTTATGTATGTTATACACTCAATACAAAAGACTGACCGG	480
QY	761	TCATACGTTAAGACGAGGGGTGTAGTAAAG	791
Db	481	TCATACGTTAAGACGAGGGGTGTAGTAAAG	511

RESULT	10
LOCUS	AV718096
DEFINITION	AV718096 FH7A Homo sapiens cDNA clone FHTABAO2 5'
ACCESSION	AV718096
VERSION	AV718096..1 GI:10815248
KEYWORDS	EST.
SOURCE	human.
	706 bp mRNA linear EST 16-OCT-2000
	clone FHTABAO2 5', mRNA sequence.

TITLE Homo sapiens CDNA FHTA clones
JOURNAL Unpublished (2000)
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@hgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES	Location/Qualifiers
source	1. .706

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="FHTABAO2"
/clone_id="FHTA"
/tissue_type="hypothalamus"
/dev_stage="Fetal"
/lab_host="BM25.8"
/note="Vector: pRR16x2; Site:1: sf11a; Site:2: sf11b1"

```

Query Match	42.6%	Score 505.4	DB 10	Length 706
Best Local Similarity	86.6%	Pred. No. 11e-99		
Matches 602	Conservative	0	Matches 38	Indels 55
				Gaps 2
OY	36	TGCATTAACCTACTACTCTACTGTGATFACATTTCAAAACCCCTGCAGATCAACAGTTATACGGTA	95	
Db	64	TGCATTAACCTACTACTCTACTGTGATFACATTTCAAAACCCCTGCAGATCAACAGTTATACGGTA	123	
OY	96	ACCAACAAGAAATGCAAGCCGTGCAATCTGACCTCTGCGCTGGGAACACCA-GCTCG	154	
Db	124	ACCAACAAGAAATGCAAGCCGACCAACAACCTCACCTCTGCGCGCTGTAAACACATGCTCG	183	
OY	155	TGCACACAGAGACTACAAAAATCAACCCAGGTCCTCTTCCCACTGCTCTACACTGCTCTGT	214	
Db	184	TGCACACAGAGACTACAAAAATCAACCCAGGTCCTCTTACCACTGCTCTACACTGCTCTGT	243	
OY	215	TTTGTGGACTATFACAAATGGCCGGGGATGAGGATTTCTTTCAAATCCGAGATAAA	274	
Db	244	TTTGTGGACTATFACAAATGGCCCTGGGGATGAGGATTTGTTCAAATCCGAGATAAA	303	
OY	275	TCAAATTTATTTATTTCTTTTAAGAACACAGTCATTTCTGATCTTCTACATGATTCGACT	334	
Db	304	TCAAATTTATTTATTTCTTTTAAGAACACAGTCATTTCTGATCTTATCATGATCTCGACT	363	
OY	335	TTTTCCATTTCAAAATTTCTTAGTAGATGCCAAACTGGGCAACGAGCACCATGAGAACTTTTGTG	394	
Db	364	TTTTCCATTTCAAAATTTCTTAGGGATGGCAACCTGGGAACAGGACACATGAGAACTTTTGTG	423	

Oy	395	TGTCAAGTTACCCGCGTAAATTTTATTTTACAAATGTAAATACAGATTTCATCTCGGGA	454
Db	424	TGTCAAGTTACCCGCGTAAATTTTATTTTACAAATGTAAATACAGATTTCATCTCGGGA	483
Oy	455	CTGATTAAGTATTCGATCGCTACCAAGAGACACACAGGCCATTTTAAACATCAACCCCAAA	514
Db	484	CTGATTAAGTATTCGATCGCTACCAAGAGACACACAGGCCATTTTAAAGCATCAACCCCA --	541
Oy	515	AATCTCTTGGGGGCTAAAGATTCTCTCTGTTCATCTGGGCAATTCATGTTCTACTCTCT	574
Db	542	-----TACTCTCT	549
Oy	575	TTGCGCTAACGATGTTCTGACCAACAGGACGCCGAGAGACACAGAATGTGAAGAAATGCTCT	634
Db	550	ATGCGCTAACGATGACCCCTGACCAACAGGACGCCGAGAGACACAGAATGTGAAGAAATGCTCA	609
Oy	635	TTCCCTTAATCGATGATGGGTCTAACTGCGGCAATGAAATGTAATTAACATCTGCTCAAGTC	694
Db	610	TTCCCTTAATCGAGACTGCTCTAATCTGCGCATGAAATAGCTAAATACATCTGNCAGATC	669
Oy	695	ATTTCCTGAGATTAAATTCCTAAATGTTATGTATG 729	
Db	670	ATTTCCTGAGATTATGTCTGAATATCCATGATG 704	

RESULT	11
LOCUS	B624828
DEFINITION	B624828 RIKEN full-length enriched, adult male colon Mus musculus cDNA clone 9030207C05 5', mRNA sequence.

VERSION	BB624828.1	GI:15398391
KEYWORDS	EST.	

SOURCE	house mouse.
ORGANISM	Mus musculus

REFERENCE
Alekawa, T., Carlinici, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.
1 (bases 1 to 674)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 674)

TITLE HIRAMOTO, K., HORII, F., ICHII, Y., ITO, M., KAWAI, J., KONNO, H., KOUDE
JOURNAL 'M., KOYA, S., MATSUYAMA, T., MIYAZAKI, A., NOMURA, K., OHNO, M.,
COMMENT OKAZAKI, Y., OKIDO, T., SAITO, R., SAKAI, C., SAKAI, K., SANO, H., SASAKI,
'D., SHIBATA, K., SHINAGAWA, A., SHIRAKI, T., SOGABE, Y., SUZUKI, H.,
Tagawa, A., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., TOYA, T.,
Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
Unpublished (2001)
Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel.: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-res@gsic.riken.go.jp,
URL: <http://genome.gsic.riken.go.jp/>
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh
, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. *Genome Res.* 10 (10), 1617-1630 (2000)
Wagci, K., Fujisake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura
, S., Kawai, U., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--334-format sequencing pipeline with 384 multicapillary sequencer. *Genome Res.* 10 (11), 1757-1771 (2000)

Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y. Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11(2): 281-289 (2001).
Yamanaka, I., Kiyosawa, H., Kondo, T., Shingawa, A., Aizawa
K., Fukuda, S., Hara, A., Itoh, M., Kawa, J., Shibata, K., Aikawa, T.,
Ishii, Y. and Hayashizaki, Y.
Mapping of 19012 mouse cDNAs on mouse chromosomes. J. Struct.
Funct. Genomics 2 pre, L72-L86 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.
e mouse tissues.
Location/Qualifiers
1. 674

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="9030207C05"
/clone_id="RIKEN full-length enriched, adult male colon"
/sex="male"
/tissue_type="colon"
/dev_stage="adult"
/lab_host="DH10B"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5':
GAGAGAGAGAGATCCACACGCTTTTTTTTTTTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 10.0 and subtraction to Rot = 185.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5' GAGAGAGAGATTCGCGATTATTAATTATTCACCCCCCCC
3']. cDNA was cloned into the XhoI and BamHI sites.
Vector: a modified pluvscript KS(+) after bulk excision
from Lambda FIC I. Cloning sites, 5' end: SalI; 3' end:

```

	BamHI ^a			
BASE COUNT	176 a	171 c	122 g	205 t
ORIGIN				

Query Match	37.98;	Score 449.4;	DB 10;	Length 674;
Best Local Similarity	84.18;	Pred. No. 1.0e-87;		
Matches 507;	Conservative 0;	Mismatches 96;	Indels 0;	Gaps 0;

[illegible]

Oy	174 TCACCCAGGCTCTTCCCACTGCTTCACACTGTCTGTTTTTTGGTAGACTATTCACAA	233
Db	132 TCACCCAGGCTCTTCCCACTGCTTCACACTGTCTGTTTTTTGGTAGACTATTCACGA	191

[illegible]

Dc	TTAAGAACACGGTCATCTGTCATCTACTAAATGAATTCAACTTTCATTAATAAATTTCTTA	311
Dd		
Df	GTGATGCCAAATCTGGGAACAAGACGACCAGTGAGAACTTTGTGTGTCAGAAGTTACCTCCGTCA	413

Db	312	GTGATGCTAAACGTGGAGCCGGGCTCTGAGAACCTTGTTGTGCGAAGTACTTCAGTCA	371
Oy	414	TATTTTATTTACAACTGATATATCAGTATTTTCATTCCTCGGGAGCTATTAACATPAGATGGCT	473

D6 372 CATTTATTATTACATGATATCATGTAATAGCTTCCGGGGTTGATACCATTTAGCCGT 431

QY 474 ACCAGAGACCACCAGGCCATTAAAAATCCAAACCCAAAATACTCTGGGGCTAGA 533

Db 432 ACCTGAAGACCACAGGSCATTTAAAACTCCAGCCAGCCAGCAATCTGTGGGTCAAGA 491

Db	198	ATCA - GTACCAACAAGAAATGACAGCGCGTGACAAACCTCACCCTGGCCCTGGGAACAC	256
QY	148	CAGTGTGTGACACAGACACACACACAAATACACCAAGTCTCTTCCACAGCTGTACACTGT	207
Db	257	CAGTGTGTGACACAGACACACACAAATACACCAAGTCTCTTCCACAGCTGTACACTGT	316
QY	208	CCTGTATTTTGTGACATTCATCACAAATGGCCGTGGCATGAGCATTTTCTTCAATTCG	267
Db	317	CCTGTATTTTGTGACATTCATCACAAATGGCCGTGGCATGAGCATTTTCTTCAATTCG	376
QY	268	GAGTAAATCAAACTTTATTTATTTTCTTAAAGAACAGTCAATTTGATCTTCATGAT	327
Db	377	GAGTAAATCAAACTTTATTTATTTTCTTAAAGAACAGTCAATTTGATCTTCATGAT	436
QY	328	TCTGACTTTTCCATTCAAAAATTTCTTAAGATGATGCCAAATGGGAACAGCACTGAGAC	387
Db	437	TCTGACTTTTCCATTCAAAAATTTCTTAAGATGATGCCAAATGGGAACAGCACTGAGAC	496
QY	388	TTTTGTGTGACAGTATACCTCCGTCAATTTTATTTTCAATATGATATACAGATTTTCAT	447
Db	497	TTTTGTGTGACAGTATACCTCCGTCAATTTTATTTTCAATATGATATACAGATTTTCAT	556
QY	448	CCTGGGACTGATTAACATAT - CGATGGCTACCAAGACACACACAG - CCATTTAAACATC	504
Db	557	CCTGGGACTGATTAACATATCGATCCCTACCAAGACACACAGGCGCATTTTAAACATC	616
QY	505	CAACCCCAAAAAA-----TCTGTGGGGGCTAAGATTTCTCTGTGTG 545	
Db	617	CAACCCCAAAAAATCTCTTTTGGGGGCTAAAAAATTTCTTCTGTG 662	
RESULT 13			
LOCUS	BG919208	877 bp	mRNA linear EST 05-JUN-2001
DEFINITION	60281803F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:4947057 5'		
ACCESSION	BG919208		
VERSION	BG919208.1	GI:14299684	
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryote: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 877)		
TITLE	NIH-MGC http://mgc.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgabs-remail.nih.gov Tissue Procurement: Jeffrey Green M.D. CDNA Library Preparation: Life Technologies, Inc. DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) Clone sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.lnl.nih.gov Plate: LHAM10896 row: e column: 10 High quality sequence stop: 815. location/Qualifiers 1. 877 /organism="Mus musculus" /strain="FVB/N" /db_xref="taxon:10090" /clone_image="IMAGE:4947057" /clone_id="NCI_CGAP_Mam6" /sex="female, virgin" /tissue_type="infiltrating ductal carcinoma" /dev_stage="5 months" /lab_host="DH10B" /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Jeffrey Green, M.D., NIH"		

ORIGIN	BASE COUNT	233 a	224 c	169 g	251 t
Query Match	35.7%;				
Best Local Similarity	82.2%;				
Matches 512:	Conservative	0;	Mismatches	108;	Indels 3; Gaps 2;
OY	114	CGGTGACAAATCTCACTCTCTGGCGTGGGAACACAGTCTGTGACACAGACTACAAA	173		
DB	242	CTTCAGCCAAATACCACTCTTCTCCCTGGGACAGACACCTGTGGCTCAAGACTACAAA	301		
OY	174	TCACCCAGAGTCTCTCCACAGCTCTACACTGCTGCTGTTTGTGGACTATACAA	233		
DB	302	TCACCCAGAGTCTCTCTCCATGTCTGTACACCGCTCTGTCTTGTGGCTATACAGA	361		
OY	234	ATGGCTGGCGATGAGATTTTCTTCAATCCGGAGTAAATCAACTTATATTTTTC	293		
DB	362	ACAGCTTGCAATGAGGATTTCTTTCAGATCCGAGTAAATCAACTCATCATTTTTC	421		
OY	294	TTAAGAACACAGTATTTCTGATCTTCTCATATTTCTGATTTCCATTCAAAATTTCTTA	353		
DB	422	TTAAGAACACAGTATTTCTGATCTTCTCATATTTCTGATTTCCATTCAAAATTTCTTA	481		
OY	354	GTGATGCCAAATGGGAACAGGACACAGAACTTTGTGTGAAGTTACTCTCGCTCA	413		
DB	482	GTGATGCCAAATGGGAACAGGACACAGAACTTTGTGTGAAGTTACTCTCGCTCA	541		
OY	414	TATTTTATTTTCAATGTTATATCAATTTTCAATTTCTGAGTGAATCAATCAATCGCT	473		
DB	542	CATTTTATTTTCAATGTTATATCAATTTTCAATTTCTGAGTGAATCAATCAATCGCT	601		
OY	474	ACCAGAACACACAGGCGCATTTTAAACATCCAAACCCAAAATCTCTTGGGGCTAGA	533		
DB	602	ACCTAAGACACACAGGCGCATTTTAAACATCCAAACCCAAAATCTCTTGGGGCTAGA	661		
OY	534	TTCCTCTGTGTCACTGAGGCGATCATGTCTTCTGCTGCTTGGCTCAATCATGATTTCTGA	593		
DB	662	TTCCTCTGTGTCACTGAGGCGATCATGTCTTCTGCTGCTTGGCTCAATCATGATTTCTGA	721		
OY	594	CCAAACAGGAGCGGAGAGACAGAAATGTGAAGAA--TGCTTTCTTCCCTTAATACAGAGTT	651		
DB	722	CCAAACAGGAGCGGAGAGACAGAAATGTGAAGAA--TGCTTTCTTCCCTTAATACAGAGTT	781		
OY	652	CGGTCTAGTCTGGCATGAATATGTAATTTACATCTGTCTCAAGTCAATTTCTGGATTAAATT	711		
DB	782	TGCTCTAGTCTGGCATGAATATGTAATTTACATCTGTCTCAAGTCAATTTCTGGATTAAATT	840		
OY	712	CTTAATGTGTAATGTATGTTATA	734		
DB	841	TTTAATGTGTAATGTGTTATA	863		
RESULT 14					
AA274112					
LOCUS	AA274112	542 bp	mRNA	linear	EST 28-MAR-1997
DEFINITION	vb92d04.r1 Soares mouse 3BDM5 Mus musculus cDNA clone IMAGE:764455				
KEYWORDS	5' similar to TR:G285995 G285995 ORF, COMPLETE CDS. ;, mRNA sequence.				
ACCESSION	AA274112				
VERSION	AA274112.1	GI:1912557			
KEYWORDS	EST.				
ORGANISM	house mouse.				
SOURCE	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Scuriognathi; Muridae; Murinae; Mus.				
AUTHORS	Maier, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, B., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R., and Waterston, R.				
TITLE	The WashU-HMNI Mouse EST Project				
JOURNAL	Unpublished (1996)				

[illegible]

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2002, 22:55:05 ; Search time 3226 Seconds
(Without alignments)
10708.317 Million cell updates/sec

Title: US-09-835-922-1

Perfect score: 1187

Sequence: 1 ggtaccggggagatcctctag.....tgtgttcagactcgttaaa 1187

Scoring table: IDENTITY_NDC

Gapop 10.0 , Gapept 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenBank:*

1: gb_ba:*

2: gb_htg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vi:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vi:*

30: em_htg_hum:*

31: em_htg_in:*

32: em_htg_inv:*

33: em_htg_mus:*

34: em_htg_pln:*

35: em_htg_rod:*

36: em_htg_man:*

37: em_htg_vtl:*

38: em_sy:*

39: em_htgo_hum:*

40: em_htgo_mus:*

41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1160	97.7	1424	9 BC017898	BC017898 Homo sapi
2	1153.4	97.2	1332	9 AF313449	AF313449 Homo sapi
3	1152	97.1	1152	9 AF321815	AF321815 Homo sapi
4	1110.4	93.5	1951	9 AB056385	AB056385 Macaca fa
5	1108.8	93.4	2266	9 AB062981	AB062981 Macaca fa
6	1097	92.4	328187	2 AC117393	AC117393 Homo sapi
7	1093.8	92.1	1143	6 AX411401	AX411401 Sequence
8	1093.8	92.1	1143	9 AF310685	AF310685 Homo sapi
9	1093.8	92.1	173270	9 AC024886	AC024886 Homo sapi
10	1027.4	86.6	1029	6 AX369349	AX369349 Sequence
11	1027.4	86.6	1029	9 AB052684	AB052684 Homo sapi
12	1025.8	86.4	1029	9 AB083596	AB083596 Homo sapi
13	732.2	61.7	2151	10 BC027381	BC027381 Mus muscu
14	732.2	61.7	2422	10 BC025428	BC025428 Mus muscu
15	732.2	61.7	211964	2 AC122038	AC122038 Mus muscu
16	729.8	61.5	1044	6 AX369354	AX369354 Sequence
17	729.8	61.5	1047	10 MM0312130	AJ312130 Mus muscu
18	701.4	59.1	1254	10 AF313450	AF313450 Rattus no
19	695	58.6	179889	2 AC096965	AC096965 Rattus no
20	466	39.3	466	6 AX369353	AX369353 Sequence
21	407.6	34.3	542	6 AX413092	AX413092 Sequence
22	307.2	25.9	1002	6 E49123	E49123 Novel G pro
23	307.2	25.9	1002	6 E50833	E50833 Novel G pro
24	307.2	25.9	1002	6 AF406692	AF406692 Homo sapi
25	307.2	25.9	1002	9 AF411113	AF411113 Homo sapi
26	307.2	25.9	1484	6 AX128499	AX128499 Sequence
27	307.2	25.9	2857	9 AF295368	AF295368 Homo sapi
28	305.6	25.7	1002	9 AB083597	AB083597 Homo sapi
29	305.6	25.7	1124	6 AR137021	AR137021 Sequence
30	305.6	25.7	1124	6 AR200987	AR200987 Sequence
31	305.6	25.7	1124	6 E36078	E36078 cDNA clone
32	303.2	25.5	1958	9 AB056816	AB056816 Macaca fa
33	302.4	25.5	1488	6 AR075038	AR075038 Sequence
34	295.4	24.9	1474	6 AR092425	AR092425 Sequence
35	263.2	22.2	356	6 AX369358	AX369358 Sequence
36	252.4	21.3	2416	6 AX329578	AX329578 Sequence
37	252.4	21.3	2416	9 HUMRSC338	D13626 Human mRNA
38	252.4	21.3	8565	9 AF411849S2	AF456925 Homo sapi
39	252.4	21.3	183861	9 AC078816	AC078816 Homo sapi
40	250.8	21.1	1027	6 AR154244	AR154244 Sequence
41	236	19.9	190799	2 AC128510	AC128510 Rattus no
42	230.8	19.4	1785	10 BC028995	BC028995 Mus muscu
43	230.8	19.4	1847	10 AF177211	AF177211 Mus muscu
44	230.8	19.4	191305	2 AC115919	AC115919 Mus muscu
45	219.2	18.5	1690	10 RN076206	U76206 Rattus norv

ALIGNMENTS

RESULT 1

BC017898

LOCUS

DEFINITION Homo sapiens, Purinergic receptor P2Y, G protein-coupled, 12, clone

ACCSSION MGC:23802 IMAGE:4251263, mRNA, complete cds.

VERSION BC017898

KEYWORDS BC017898.1 GI:17389766

SOURCE MGC.

ORGANISM Homo sapiens.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 1424)

Strausberg,R.

TITLE Direct Submission

JOURNAL

Submitted (03-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550, USA

REMARK

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: CLONTECH

CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNI)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.shgc.stanford.edu>
Contact: (Dickson, Mark) mdickson@stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNI at: <http://image.llnl.gov>
Series: IRAL Plate: 35 Row: e Column: 21
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA g1: 1232482.

FEATURES

source

1..1424
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:23802 IMAGE:4251263"
/tissue_type="Prostate"
/clone_lib="NIH_MGC_83"
/lab_host="DH108"
/note="Vector: pDNR-LIB"
245..1273
/codon_start=1
/product="Purkinergic receptor P2Y, G protein-coupled, 12"
/protein_id="AAH17898.1"
/db_xref="GI:17389767"
/translation="MOAVDLNLSAPNTSLCTPDKYKIQVLPILVTLFVGLITNG
LAMRIFFOIRSKSNFIIFLKNTVSDILMTPEPRLISDAGTSGPLTPVQCQVSY
IFVTYVISTSLGLTIDRYOKTRTPFTSNKRLILGKILISVYMAWFLSLPDM
ILNRPDRKNKSCFLSEFGLWHELVNVCIVTEWLNFLIVACVTLITKELYS
SYVTRGVGKVRKVNVRVFIILAVFELCFVPHFARLIPVLSQTRVDVDCAEML
FYVEETLWLTSLNCLDPLFIYFLCKSPRNSLISMLKPNSTSLSDNRKKEQGG
DPNEETPM"

CDS

BASE COUNT 465 a 309 c 244 g 406 t
ORIGIN
Query Match 97.7%; Score 1160; DB 9; Length 1424;
Best Local Similarity 100.0%; Pred. No. 2e-274;
Matches 1160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

28 CTGCAGGCTGCATTAATCTACTACTGATGATTCATTCAACCTCCAGAAATGACAGTT 87
166 CTGCAGGCTGCATTAATCTACTACTGATGATTCATTCAACCTCCAGAAATGACAGTT 225
88 ATCAGGTAACCAAGAAATGCAAGCCGTCGACAAATCTACCTGCGCTGGAAACAC 147
226 ATCAGGTAACCAAGAAATGCAAGCCGTCGACAAATCTACCTGCGCTGGAAACAC 285
148 CAGTCTGTGCACGACGACTACAAATTCACCCAGGTCCTCTCCACTGCTTACACTGT 207
286 CAGTCTGTGCACGACGACTACAAATTCACCCAGGTCCTCTCCACTGCTTACACTGT 345
208 CCGTGTCTTTTGGACTATATCAAAATGCGCGGAGATGAGATTTCTTTAAATCCG 267
346 CCGTGTCTTTTGGACTATATCAAAATGCGCGGAGATGAGATTTCTTTAAATCCG 405
268 GAGTAATCAAACTTATATTTTCTTAAGAACACAGCATTTGATCTTCATGTAT 327
406 GAGTAATCAAACTTATATTTTCTTAAGAACACAGCATTTGATCTTCATGTAT 465
328 TGTGATTTTCCATTCAAAATCTTATGATGCAACTGGAAACGACCACTGGAAC 387

Db 466 TGTGACTTTTCCATTCAAAATTTCTTAGTGATGCCAAACTGGGAACAGGACCACTGGAAC 525
Qy 388 TTTTGTGTGCAGATTAATCTCCGTCATATTTTATTTACAAATGATATACATTTTCATT 447
Db 526 TTTTGTGTGCAGATTAATCTCCGTCATATTTTATTTACAAATGATATACATTTTCATT 585
Qy 448 CTTGGGACTGATTAATCTATGATGCTTCCAGAGACCAAGCCACTTTAAACATCCAA 507
Db 586 CTTGGGACTGATTAATCTATGATGCTTCCAGAGACCAAGCCACTTTAAACATCCAA 645
Qy 508 CCCCCAAAATCTCTGGGGGCTAAGATTTCTCTGTCATCTGGGCAATTCATGCTCT 567
Db 646 CCCCCAAAATCTCTGGGGGCTAAGATTTCTCTGTCATCTGGGCAATTCATGCTCT 705
Qy 568 ACTCTCTTGGCTTAACATGATTTCTGACCAACAGGACCCAGAGACAAAGATGGAAGAA 627
Db 706 ACTCTCTTGGCTTAACATGATTTCTGACCAACAGGACCCAGAGACAAAGATGGAAGAA 765
Qy 628 ATGCTCTTCTTAAATCAAGATTCGGTCTAGTCTGATGGAATTAATTAATTCATCTG 687
Db 766 ATGCTCTTCTTAAATCAAGATTCGGTCTAGTCTGATGGAATTAATTAATTCATCTG 825
Qy 688 TCAAGTCAATTTCTGATTAATTTCTTAATTTGTTATGTTATGTTATACATCAATACAA 747
Db 826 TCAAGTCAATTTCTGATTAATTTCTTAATTTGTTATGTTATGTTATACATCAATACAA 885
Qy 748 AGAATCTACCGGTCATACGTAAGAACAGGAGGTGTAAGTAAAGTCCCAAGAAAAGT 807
Db 886 AGAATCTACCGGTCATACGTAAGAACAGGAGGTGTAAGTAAAGTCCCAAGAAAAGT 945
Qy 808 GAAGCTCAAGTTTCTCATTAATTCATGCTGATTTCTTTATTTGTTTGTCTCTTCCATTT 867
Db 946 GAAGCTCAAGTTTCTCATTAATTCATGCTGATTTCTTTATTTGTTTGTCTCTTCCATTT 1005
Qy 868 TGCCCGAATTCCTTACACCTGAGCCAAACCCGGATGCTTTGACGCGCATGCTGAAA 927
Db 1006 TGCCCGAATTCCTTACACCTGAGCCAAACCCGGATGCTTTGACGCGCATGCTGAAA 1065
Qy 928 TACTCTGTTATATGGAAGAGAGACACTGTGTGTTAACTTCTTAATGATGCTGTGA 987
Db 1066 TACTCTGTTATATGGAAGAGAGACACTGTGTGTTAACTTCTTAATGATGCTGTGA 1125
Qy 988 TCGGTCATCATTTTCTTCTTGGCAAGTCTTCAGAAATTCCTGTATAGTATGCTGAA 1047
Db 1126 TCGGTCATCATTTTCTTCTTGGCAAGTCTTCAGAAATTCCTGTATAGTATGCTGAA 1185
Qy 1048 GTGCCCAATTCGCAACATCTCTGCCAGGACCAATGAGAAAAGAACAGATGTGG 1107
Db 1186 GTGCCCAATTCGCAACATCTCTGCCAGGACCAATGAGAAAAGAACAGATGTGG 1245
Qy 1108 TGACCCAAATGGAAGAGACTGCAATGTAAACAAATTAAGTAAAGAAATTTCAATCTTT 1167
Db 1246 TGACCCAAATGGAAGAGACTGCAATGTAAACAAATTAAGTAAAGAAATTTCAATCTTT 1305
Qy 1168 TGTGTCAGAACTGTTAAA 1187
Db 1306 TGTGTCAGAACTGTTAAA 1325

RESULT 2
AF313449 1332 bp mRNA linear PRI 12-JAN-2001
LOCUS
DEFINITION
AF313449 Homo sapiens P2Y12 platelet ADP receptor mRNA, complete cds.
VERSION
AF313449.1 GI:12083901
KEYWORDS
SOURCE
ORGANISM
Homo sapiens.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1332)
Hollinger, G., Jantzen, H.-M., Vincent, D., Li, G., England, L.,
Ramakrishnan, V., Yang, R.-B., Nurdien, P., Nurdien, A., Julius, D. J. and

[illegible]

OY	748	AGAACTGACGGGCATCATGTGAAGAAGAGGGGTCAGTAAAGCCCAAGAAAAAAGT	807
Db	889	AGAACTGACGGGCATCATGTGAAGAAGAGGGGTCAGTAAAGTCCCCAGAAAAAAGT	948
OY	808	GAACTCAAAGTTTTCATTATCATATGCGTAGTAATTCCTTAATTTGTTTTGTCCTTCCATT	867
Db	949	GAACTCAAAGTTTTCATTATCATATGCGTAGTAATTCCTTAATTTGTTTTGTCCTTCCATT	1008
OY	868	TGCCCGAATTCCTTACACCCCTGAGCCAAAACC GGAGTGCTTTTGACTGACACTGCTGAAA	927
Db	1009	TGCCCGAATTCCTTACACCCCTGAGCCAAAACC GGAGTGCTTTTGACTGACACTGCTGAAA	1068
OY	928	TACTCTGTTCTTAATGGAAGAAGAGACACCTGCTGGTTACTCTCCTTAATTCGATGCTGGA	987
Db	1069	TACTCTGTTCTTAATGGAAGAAGAGACACCTGCTGGTTACTCTCCTTAATTCGATGCTGGA	1128
OY	988	TCCGTTTCATCATATTTTTCCTTTGCAAGTCCTTCAGAAATTCCTTGATAATGATGCTGAA	1047
Db	1129	TCCGTTTCATCATATTTTTCCTTTGCAAGTCCTTCAGAAATTCCTTGATAATGATGCTGAA	1188
OY	1048	GTGCCCCAATTCCTGACACATCTCTGTCCCAGACAAATGGAAGAAAAAGACAGATGTGG	1107
Db	1189	GTGCCCCAATTCCTGACACATCTCTGTCCCAGACAAATGGAAGAAAAAGACAGATGTGG	1248
OY	1108	TGACCCAATGGAAGAGATCCCAATGTAACAATTAACTAAGAAATATTTCAATCTCTT	1167
Db	1249	TGACCCAATGGAAGAGATCCCAATGTAACAATTAACTAAGAAATATTTCAATCTCTT	1308
OY	1168	TGTGTTTCAGAACTGCTTAAA	1187
Db	1309	TGCGTTTCAGAACTCATTTAA	1328
RESULT 5			
AB062981			
LOCUS			
DEFINITION	AB062981	2266 bp mRNA linear PRI 13-JUN-2001	
ACCESSION		Macaque fascicularis brain cDNA clone:QmoA-10634, full insert	
VERSION	AB062981.1	GI:14388396	
KEYWORDS		oligo capping; fts (full insert sequence).	
SOURCE		Macaque fascicularis adult male medulla oblongata cDNA to mRNA,	
ORGANISM		clone_1fb:macaque brain cDNA library QmoA clone:QmoA-10634.	
		Eukarya; Metazoa; Chordata; Granata; Vertebrata; Euteleostomi;	
		Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;	
		Cercopithecinae; Macaca.	
REFERENCE		1 (sites)	
AUTHORS		Osada,N., Hida,M., Kusuda,J., Tanuma,R., Iseki,K., Hirai,M.,	
TITLE		Terao,K., Suzuki,Y., Sugano,S. and Hashimoto,K.	
JOURNAL		Isolation of full-length cDNA clones from macaque brain cDNA	
AUTHORS		libraries	
TITLE		Unpublished	
JOURNAL		2 (bases 1 to 2266)	
AUTHORS		Hashimoto,K., Osada,N., Hida,M., Kusuda,J. and Sugano,S.	
TITLE		Direct Submission	
JOURNAL		Submitted (11-JUN-2001) Katsuyuki Hashimoto, National Institute of	
		Infectious Diseases, Division of Genetic Resources; 23-1, Toyama	
		1-chome, Shinjuku-Ku, Tokyo 162-8640, Japan	
		(E-mail:Khashl@nih.go.jp, URL:http://www.nih.go.jp/yoken/genebank/,	
		Tel.:81-3-5285-1111(ex 2120), Fax:81-3-5285-1181)	
COMMENT		Lab host: TOP10	
		Vector: pME18S-F13 (Acc.No. AB009864)	
		R. Site1: DraIII (CACCTGTC)	
		R. Site2: DraIII (CACCTGTC)	
		Description: 1st strand cDNA was primed with an oligo(dt) ²⁷ primer	
		[AAGTGGCCCTTTTATTTTTTTTTT]; double-stranded cDNA was synthesized	
		using specific 5' and 3' primers and amplified by PCR. The PCR	
		product was digested with SfiI and size selection was performed to	
		exclude fragments <1.5kb.The SfiI-digested PCR product was cloned	
		into distinct DraIII sites of pME18S-F13. XhoI sites just outside	
		the DraIII sites can be used to isolate the cDNA insert. Libraries	

were constructed by oligo-capping method (Sugano et al., Institute of Medical Science, University of Tokyo).
Custom primer used for sequencing (5' end primer (CTGTGCTCTAAAGCTGG): 3' end primer (CGACCTGAGCTCGACGACAC)).
Location/Qualifiers

FEATURES

Source

1..2266
/organism="Macaca fascicularis"
/db_xref="taxon:9541"
/clone="OmoA-10634"
/sex="male"
/tissue_type="medulla oblongata"
/clone_lib="macaque brain cDNA library OmoA"
/dev_stage="adult"
240..1268
/codon_start=1
/product="hypothetical protein"
/protein_id="BAB60747.1"
/db_xref="GI:14388397"
/translation="MQAIDNLVSAPNTSLCTRDYKITQVLFPLLYTLFVGLITNS
LAMRIFFOIRKSNFIIFLKNVSDILMILTEPFLSPAKIAGAPLEFVQVTSV
IEFYMYISISGLTIDRYOKTRPFSNPKNLGLILSVLIWAFELISLNM
ILNRRPDKNNVKCSFLKSEGLVMHEIYNTYCOVIFPNIPLIVTCYTLIRKELR
SVYRPGVGRKPRKRVNYVFIIIVAFICFPFPHARIPYITISQTRDPDCAENTL
FYKESTLMLISLNACLDPIYIFLCKSPRNLISMLKCPNSATISOSODNRKEDQGG
DPNEETPM"

CDS

BASE COUNT 794 a 431 c 349 g 692 t
ORIGIN

Query Match

best Local Similarity 97.28; Score 1108.8; DB 9; Length 2266;
Pred. No. 7; 7e-262;

Matches 1128; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 28 CTGCGAGCTGCAATTAATCTACTACTGATGATACATTCAAACCTCCAGATCAACAGTT 87
DB 161 CTGACGCTTGCAATTAATCTACTACTGATGATACATTCAAACCTCCAGATCAACAGTT 220
QY 88 ATCAGGTACCAACAGAAATGCAAGCGGTGACATTCACCTGCGCGTGGAAAC 147
DB 221 ATCAGGTACCAACAGAAATGCAAGCGGTGACATTCACCTGCGCGTGGAAAC 280
QY 148 CAGCTGTCGACAGACAGACAGAAATGCAAGCGGTGACATTCACCTGCGCGTGGAAAC 207
DB 281 CAGCTGTCGACAGACAGACAGAAATGCAAGCGGTGACATTCACCTGCGCGTGGAAAC 340
QY 208 CCGTGTGTTGTTGACATTCACAAATGCGCGGTGACATTCACCTGCGCGTGGAAAC 267
DB 341 CCGTGTGTTGTTGACATTCACAAATGCGCGGTGACATTCACCTGCGCGTGGAAAC 400
QY 268 GAGTAATCAAACTTTATTTTCTTAAGAACACAGTATTCATCTTCATGAT 327
DB 401 GAGTAATCAAACTTTATTTTCTTAAGAACACAGTATTCATCTTCATGAT 460
QY 328 TGTGACTTTCATCAAAATTTCTAGTATGCAAACTGGAAACAGACACCTGAGAAC 387
DB 461 TGTGACTTTCATCAAAATTTCTAGTATGCAAACTGGAAACAGACACCTGAGAAC 520
QY 388 TTTTGTGTCAGATTCCTCCGTCATATTTTATTCACATGATATTCATGATTTTCTT 447
DB 521 TTTTGTGTCAGATTCCTCCGTCATATTTTATTCACATGATATTCATGATTTTCTT 580
QY 448 CTGCGAGCTGATACATTCATGATGCTACAGAAAGACAGCGGCACTTTAAACATCAA 507
DB 581 CTGCGAGCTGATACATTCATGATGCTACAGAAAGACAGCGGCACTTTAAACATCAA 640
QY 508 CCCCCAAAATCTCTGGGGGCTAAGATTCCTCTGTTGTCATCTGGGCACTTCACTGTTCTT 567
DB 641 CCCCCAAAATCTCTGGGGGCTAAGATTCCTCTGTTGTCATCTGGGCACTTCACTGTTCTT 700
QY 568 ACTCTCTTCTGCTACATGATTCGACCAAGACAGGACAGAGACAAAGATGGAAGAA 627
DB 701 ACTCTCTTCTGCTACATGATTCGACCAAGACAGGACAGAGACAAAGATGGAAGAA 760

QY 628 ATGCTCTTCTCTTAATACAGATTCGCTAGTCTGCGACAAATATTAATTCATCTG 687
DB 761 ATGCTCTTCTCTTAATACAGATTCGCTAGTCTGCGACAAATATTAATTCATCTG 820
QY 688 TCAAGTATTTTCTGATTAATTTCTTAATTTGTTATTTGTTATTAACCTCATTACAA 747
DB 821 TCAAGTATTTTCTGATTAATTTCTTAATTTGTTATTTGTTATTTGTTATTTGTTATTT 880
QY 748 AGAAGCTGACGGCTATACGTAAGAACAGAGGGGTGTAAGAGTCCCGAGGAAAAGGT 807
DB 881 AGAAGCTGACGGCTATACGTAAGAACAGAGGGGTGTAAGAGTCCCGAGGAAAAGGT 940
QY 808 GAAAGTCAAAATTTCTATTAATTCATTCGCTGATTTCTTAATTTGTTATTTGTTATTTGTT 867
DB 941 GAAAGTCAAAATTTCTATTAATTCATTCGCTGATTTCTTAATTTGTTATTTGTTATTTGTT 1000
QY 868 TGCCCGAATTCCTTACACCTCGAGCCAAACCCGGAGATGTTTACATGCTGCTGAAA 927
DB 1001 TGCCCGAATTCCTTATACCTCGAGCCAAACCCGGAGATGTTTACATGCTGCTGAAA 1060
QY 928 TACTCTGTTATGTAAGAGAGACACTGTGTGTTAATCTTAAATGATGATGCTGGA 987
DB 1061 TACTCTGTTATGTAAGAGAGACTGTGTGTTAATCTTAAATGATGATGCTGGA 1120
QY 988 TCCGTTATCTATTTTCTTCTTGGCAAGTCTTCAGAAATCTCTGATTAATGATGCTGGA 1047
DB 1121 TCCGTTATCTATTTTCTTCTTGGCAAGTCTTCAGAAATCTCTGATTAATGATGCTGGA 1180
QY 1048 GTGCCCCAATTTGCAACATCTCTGCTCCAGACATAGAGAAAAGACAGAGTGTG 1107
DB 1181 GTGCCCCAATTTGCAACATCTCTGCTCCAGACATAGAGAAAAGACAGAGTGTG 1240
QY 1108 TGACCCCAATGAGAGACTCCATGTAACAAATTAAGAGAAATATTCATCTCT 1167
DB 1241 TGACCCCAATGAGAGACTCCATGTAACAAATTAAGAGAAATATTCATCTCT 1300
QY 1168 TGTGTTGAGACTGTTAAA 1187
DB 1301 TGTGTTGAGACTGTTAAA 1320

RESULT 6
AC117393/c 328187 bp DNA linear HTG 29-APR-2002
LOCUS
DEFINITION
Homo sapiens chromosome 3 clone Rpl1-3729, *** SEQUENCING IN
PROGRESS ***, 17 unordered pieces.
AC117393 AC012057
AC117393.2 GT:20279204
VERSION
AC117393.2
KEYWORDS
HTG; HTGS_PHASE1; HTGS_ACTIVEFIN.
SOURCE
Homo sapiens.
ORGANISM
Homo sapiens

REFERENCE
AUTHORS
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 328187)
Muzny,D.M., Adams,C., Adio-Oduola,B., All-ouman,F.R., Allen,C.,
Alshrooks,S.L., Benton,J., Bimberg,K., Blankenburg,K., Banks,T.,
Barberia,J., Bowles,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Butrell,K.L., Byrd,N.C.,
Carton,T.F., Carter,R., Cavazos,S.R., Chacko,J., Chaves,D.,
Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durkin,K.J.,
Earmhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escoto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorelli,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hoggins,M., Holloway,C., Hollins,B.,
Homsl,F., Howard,S., Huber,J., Huiyk,S., Hume,J., Jackson,L.E.,

Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
 Karlsson, E., Kelly, S., Khan, V., King, L., Korvah, J., Kovar, C.,
 Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.,
 Li, J., Li, Z., Licharge, O., Lieu, J., Liu, W., Lloisged, H.,
 Lozano, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
 Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
 Massey, E., Mawhney, E., McLeod, M. P., Meador, M., Mel, G., Metker, M.,
 Miner, G., Miner, Z., Mitchell, T., Mohabat, K., Morgan, M., Morris, S.,
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
 Nguyen, N., Nickerson, E., Nwokkwo, S., Ogub, M., Okunolu, G.,
 Orangun, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
 Peters, L., Pickens, R., Primus, E., Pu, L. L., Qulles, M., Ren, Y.,
 Rives, M., Rojas, A., Rojuben, I., Rolfe, M., Ruiz, S., Savary, G.,
 Scherer, S., Scott, G., Shen, H., Shoshitari, N., Sisson, I.,
 Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H.,
 Sutton, A., Sytek, A., Taber, P., Tameris, A., Tameris, K., Tang, H.,
 Tasey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
 Umani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
 Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K.,
 Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstock, G. and Gibbs, R.

Unpublished
 Direct Submission
 2 (bases 1 to 328187)
 Worley, K. C.

Direct Submission
 Submitted (10-APR-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 328187)
 Worley, K. C.

Direct Submission
 Submitted (29-APR-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On or before Apr 26, 2002 this sequence version replaced
 gi:8072477, gi:20127689.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: HDGP
 Center clone name: RP11-372J9
 ----- Summary Statistics
 Sequencing vector: Plasmid
 Chemistry: Dye-terminator Big Dye 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 345031 bases at least Q40
 Consensus quality: 361075 bases at least Q30
 Consensus quality: 371672 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 17 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 2010: contig of 2010 bp in length
 * 2011 2110: gap of unknown length
 * 2111 4550: contig of 2440 bp in length
 * 4551 4650: gap of unknown length
 * 4651 7410: contig of 2660 bp in length
 * 7411 7410: gap of unknown length
 * 7411 10060: contig of 2650 bp in length
 * 10061 10160: gap of unknown length
 * 10161 14011: contig of 3851 bp in length

14012 14111: gap of unknown length
 * 14112 18735: contig of 4624 bp in length
 * 18736 18836: gap of unknown length
 * 18836 23974: contig of 5138 bp in length
 * 23974 24074: gap of unknown length
 * 24074 30069: contig of 5996 bp in length
 * 30069 30170: gap of unknown length
 * 30170 37639: contig of 7470 bp in length
 * 37640 37739: gap of unknown length
 * 37740 46044: contig of 8305 bp in length
 * 46045 46145: gap of unknown length
 * 46145 57813: contig of 11659 bp in length
 * 57814 57913: gap of unknown length
 * 57914 73711: contig of 15798 bp in length
 * 73712 73812: gap of unknown length
 * 73812 89444: contig of 15633 bp in length
 * 89445 89544: gap of unknown length
 * 89545 110717: contig of 21173 bp in length
 * 110718 110817: gap of unknown length
 * 110818 131687: contig of 20870 bp in length
 * 131688 131787: gap of unknown length
 * 131788 165404: contig of 33617 bp in length
 * 165405 165505: gap of unknown length
 * 165505 328187: contig of 162683 bp in length.

FEATURES
 source
 1. 328187
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="3"
 /clone="RP11-372J9"

BASE COUNT 98046 a 61482 c 63090 g 103950 t 1619 others
 ORIGIN

Query Match 92.4% Score 1097; DB 2: Length 328187;
 Best Local Similarity 100.0%; Pred. No. 5.7e-259;
 Matches 1097; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 91 AGGTACCAACAAAGAAAGCAAGCCGTGACATATCTGACCTGGGACACACG 150
 Db 268734 AGGTACCAACAAAGAAAGCAAGCCGTGACATATCTGACCTGGGACACACG 268675

OY 151 TCTGTGACACAGAGACTACAAATATCCAGAGCTCTTCCACTGCTCTACACTGCTCT 210
 Db 268674 TCTGTGACACAGAGACTACAAATATCCAGAGCTCTTCCACTGCTCTACACTGCTCT 268615

OY 211 GTTTTGTGGACTTATCACAAATGCGCTGGCATGAGATTTCTTCAATCCGAG 270
 Db 268614 GTTTTGTGGACTTATCACAAATGCGCTGGCATGAGATTTCTTCAATCCGAG 268555

OY 271 TAAATCAAACTTATATTTTCTTAAGACACACTATTTCTGATCTTCATGATTTCT 330
 Db 268554 TAAATCAAACTTATATTTTCTTAAGACACACTATTTCTGATCTTCATGATTTCT 268495

OY 331 GACTTTTCATTCAAAATTTCTAGATGCCAAACTGGGACAGACACACTGAGACTTT 390
 Db 268494 GACTTTTCATTCAAAATTTCTAGATGCCAAACTGGGACAGACACACTGAGACTTT 268435

OY 391 TGTGTGCAAGTATACCTCGCATATTTTATTTTATTTTCAATGATATTCAGATTTCTTCT 450
 Db 268434 TGTGTGCAAGTATACCTCGCATATTTTATTTTATTTTCAATGATATTCAGATTTCTTCT 268375

OY 451 GGGAGCTAATCATATGATGCTGACAGAAAGCCACAGAGCCATTTAAATCCACACC 510
 Db 268374 GGGAGCTAATCATATGATGCTGACAGAAAGCCACAGAGCCATTTAAATCCACACC 268315

OY 511 CAAAATCTCTTTGGGGCTAGATTTCTCTGTGTGATCTGGGACTTCATGTTTACT 570
 Db 268314 CAAAATCTCTTTGGGGCTAGATTTCTCTGTGTGATCTGGGACTTCATGTTTACT 268255

OY 571 CTCTTGGCTTAACAGATTTGACCAACAGAGCCGAGAGACAGATGGAAGAATG 630
 Db 268254 CTCTTGGCTTAACAGATTTGACCAACAGAGCCGAGAGACAGATGGAAGAATG 268195

Db 996 CCCCAATTCGCAACATCTCTGTCGCCAGACATAGCAAAAAGACAGATGGTGTGA 1055
QY 1111 CCCAATGAAGAGACTCCCAATGTAAACAATTAAGCAAAATATTTCAATCTTTGT 1170
Db 1056 CCCAATGAAGAGACTCCCAATGTAAACAATTAAGCAAAATATTTCAATCTTTGT 1115
QY 1171 GTTCAGACTCGTTAA 1187
Db 1116 GTTCAGACTCGTTAA 1132

RESULT 8
AF310685 1143 bp DNA linear PRI 17-NOV-2001
LOCUS Homo sapiens ADP-glucose receptor gene, complete cds.
DEFINITION AF310685
ACCESSION AF310685
VERSION AF310685.1 GI:16973448
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1143)
Reinscheid, R.K., Nothacker, H.-P., Wang, Z., Zeng, J., Ehler, F.J. and
Civelli, O.
ADP-glucose activates a G-protein coupled receptor and inhibits
smooth muscle contractions
Unpublished
2 (bases 1 to 1143)
Nothacker, H.-P., Reinscheid, R.K., Wang, Z. and Civelli, O.
Direct Submission
Submitted (03-OCT-2000) Pharmacology, UC Irvine, 354 Med Surge II,
Irvine, CA 92697, USA
location/Qualifiers
1. 1143
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"
/map="3q21-q25"
/product="ADP-glucose receptor"
/note="ADP-R: G protein-coupled receptor"
/codon_start=1
/product="ADP-glucose receptor"
/protein_id="AL32292.1"
/db_xref="GI:16973449"
/translation="MQAVNDLTSAPGNTSICTRDYKITIYVLEPLTYLVEVGLING
LAKRIFQIRSKSNFIIFLKNYISDLMLITPEKILSDAKLGTGRLTFVCOVTSV
IEFVWISISFLGILITIDRYOKTTRPFKTSNPKNLGAILSVIWAFFILSLPDM
ILNRQPRDKNVKCSFLKSEFLVWHEIVNYICQVIFWNLIVICVYLITRELYR
SVYRTGVGVPRKRVNRYFIIIAVFICFVPHFARIPIYTSORVDYDCAEANTL
FYVKESTLMTLSLNACLDPIYFELFKSPFNLSISMKCNKSNATSLSDNRKKEQDGG
DPERETPM"

BASE COUNT 326 a 257 c 201 g 359 t
ORIGIN

Query Match 92.1%; Score 1093.8; DB 9; Length 1143;
Best Local Similarity 99.8%; Pred. No. 3.8e-258;
Matches 1095; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 91 AGTAACCAACAGAAATGCAAGCGGTGACATCTGACCTGCGGCTGGGAACACAG 150
Db 36 AGTAACCAACAGAAATGCAAGCGGTGACATCTGACCTGCGGCTGGTAACACAG 95
QY 151 TCTGTGACACAGACATCAACAAATCACCAGGCTCTCCCACTGCTACACTGTCT 210
Db 96 TCTGTGACACAGACATCAACAAATCACCAGGCTCTCTCCCACTGCTACACTGTCT 155
QY 211 GTTTTGTGACATTCACAAATGCGCTGGCATGAGATTTCTTTCAATCCGAG 270
Db 156 GTTTTGTGACATTCACAAATGCGCTGGCATGAGATTTCTTTCAATCCGAG 215

QY 271 TAAATCAACTTTATATTTTCTTAAGAACACATCATTTCTGATCTCTCAAGTCT 330
Db 216 TAAATCAACTTTATATTTTCTTAAGAACACATCATTTCTGATCTCTCAAGTCT 275
QY 331 GACTTTCATCAAAATTCCTAGTATGCCAACTGGGAACAGACACATGAACTTT 390
Db 276 GACTTTCATCAAAATTCCTAGTATGCCAACTGGGAACAGACACATGAACTTT 335
QY 391 TGTGTCAAGTTCACCTCCGTCATATTTTATTTTCAATGATATCATTTCTCT 450
Db 336 TGTGTCAAGTTCACCTCCGTCATATTTTATTTTCAATGATATCATTTCTCT 395
QY 451 GGCATGATTAATCATGATGCTACCAAGAACACACAGGCCATTTAAACATCAACC 510
Db 396 GGCATGATTAATCATGATGCTACCAAGAACACACAGGCCATTTAAACATCAACC 455
QY 511 CAAAATCTCTGGGGCTAAGATTCCTCTGTTGTCATCTGGGCATTCATGTTCTACT 570
Db 456 CAAAATCTCTGGGGCTAAGATTCCTCTGTTGTCATCTGGGCATTCATGTTCTACT 515
QY 571 CTCTTTCCTTAACATGATTCGACCAACAGGACCGACAGACAAAGATGGAAGAAATG 630
Db 516 CTCTTTCCTTAACATGATTCGACCAACAGGACCGACAGACAAAGATGGAAGAAATG 575
QY 631 CTCTTTCCTTAACATGATTCGACCAACAGGACCGACAGACAAAGATGGAAGAAATG 690
Db 576 CTCTTTCCTTAACATGATTCGACCAACAGGACCGACAGACAAAGATGGAAGAAATG 635
QY 691 AGTATTTTTCGGAATTAATTTCTTAATTTGATTTGATTTATCATCTATTAACAAAGA 750
Db 636 AGTATTTTTCGGAATTAATTTCTTAATTTGATTTGATTTATCATCTATTAACAAAGA 695
QY 751 ACTGTACCGGTCATACGATTAAGAACAGAGGGGTGTAGTAAGTCCCGAGAAAAGTGA 810
Db 696 ACTGTACCGGTCATACGATTAAGAACAGAGGGGTGTAGTAAGTCCCGAGAAAAGTGA 755
QY 811 CGTCAAGTTTTCATATCATGCTGATATCTTTATTTGTTTGTCTTTCATTTTGC 870
Db 756 CGTCAAGTTTTCATATCATGCTGATATCTTTATTTGTTTGTCTTTCATTTTGC 815
QY 871 CCGAATTCCTTACACCGTGGAGCCAAACCGGAGATCTTGACGCACTGCTGAAATAC 930
Db 816 CCGAATTCCTTACACCGTGGAGCCAAACCGGAGATCTTGACGCACTGCTGAAATAC 875
QY 931 TCTGTTCATGTAAGAGACACTGTGTGATTAATCTCTTAATGATGCTGAAGTG 1050
Db 876 TCTGTTCATGTAAGAGACACTGTGTGATTAATCTCTTAATGATGCTGAAGTG 935
QY 991 GTTCATCTATTTTCTCTTTCGAAGTCTTCAGAAATTCCTGATTAAGTATGCTGAAGTG 1050
Db 936 GTTCATCTATTTTCTCTTTCGAAGTCTTCAGAAATTCCTGATTAAGTATGCTGAAGTG 995
QY 1051 CCCCAATTCGCAACATCTCTGTCGCCAGACATAGCAAAAAGACAGATGGTGTGA 1110
Db 996 CCCCAATTCGCAACATCTCTGTCGCCAGACATAGCAAAAAGACAGATGGTGTGA 1055
QY 1111 CCCAATGAAGAGACTCCCAATGTAAACAATTAAGCAAAATATTTCAATCTTTGT 1170
Db 1056 CCCAATGAAGAGACTCCCAATGTAAACAATTAAGCAAAATATTTCAATCTTTGT 1115
QY 1171 GTTCAGACTCGTTAA 1187
Db 1116 GTTCAGACTCGTTAA 1132

RESULT 9
AC024886 173270 bp DNA linear PRI 28-MAR-2002
LOCUS Homo sapiens 3 BAC Rpl1-25K24 (Roswell Park Cancer Institute Human
BAC library) complete sequence.
DEFINITION AC024886
ACCESSION AC024886
VERSION AC024886.20 GI:16572827
KEYWORDS HTG.

SOURCE Homo sapiens.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 173270)
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., All-ouman,F.R., Allen,C., Alshrocks,S.L., Amaratunga,H.C., Are,J.R., Banks,T., Barbieri,J., Benton,J., Bimarge,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowe,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burich,P., Burkett,C., Burrell,K.L., Byrd,N.C., Cartron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy,Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Demn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garrett,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hatl,M., Havlak,P., Hawes,A., He,X., Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,D.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louisedge,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Meli,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabat,K., Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokwenkwo,S., Oguh,M., Okwuonon,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshitari,N., Sisson,I., Sodergren,E., Sotake,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Szatk,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalob,D., Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wlarczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Naylor,S.L., Weinstein,G. and Gibbs,R.

TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 173270)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (02-MAR-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 3 (bases 1 to 173270)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (01-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 4 (bases 1 to 173270)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (28-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT On Nov 1, 2001 this sequence version replaced gi:16515437.
INFORMATION: http://www.hgsc.bcm.tmc.edu/or_email/gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones.

Overlapping clones are noted at the beginning and end of the features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts. Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences. Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as low coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8086/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT.

FEATURES	QUALSTAT-REPORT
source	Location/Qualifiers
1. 173270	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="3"
	/clone="RP11-25K24"
	/complement(259..684)
	/rpt_family="MER57A"
	1209..1372
	/rpt_family="L2"
	complement(2596..2721)
	/rpt_family="Alusq/x"
	2722..2897
	/rpt_family="Alusp"
	2980..3091
	/rpt_family="MIR"
	3781..3979
	/rpt_family="MIR"
	complement(4002..4186)
	/rpt_family="MER5A"
	4476..4950
	/rpt_family="L1MA9"
	7145..7190
	/rpt_family="CATATA)n"
	8285..8320
	/rpt_family="CATATA)n"
	8380..8400
	/rpt_family="AT-rich"
	complement(8489..8520)
	/rpt_family="L2"
	complement(8642..8888)
	/rpt_family="MIR"
	complement(8972..9105)
	/rpt_family="L2"
	10672..11018
	/rpt_family="L1MB7"
	11019..11332
	/rpt_family="Alusp"
	11333..11418
	/rpt_family="L1MB7"
	complement(11419..11707)
	/rpt_family="Alusx"

```
repeat_region 11708..11761
/rpt_family="L1MB7"
complement(11847..12113)
repeat_region /rpt_family="L1MC5"
12114..12425
/rpt_family="Alusx"
repeat_region /rpt_family="MER5B"
13185..13306
/rpt_family="MER5B"
13767..13843
/rpt_family="(CATATA)n"
repeat_region 13844..13985
/rpt_family="Tiger3(Golem)"
13983..14213
/rpt_family="Tiger3(Golem)"
repeat_region complement(14962..15308)
/rpt_family="L2"
repeat_region complement(15513..15816)
/rpt_family="Alusx"
16407..16634
/rpt_family="L2"
repeat_region complement(16868..16995)
/rpt_family="L2"
21821..22127
/rpt_family="Alusq"
repeat_region complement(22172..22439)
/rpt_family="Aluo"
23743..24050
/rpt_family="Alusx"
repeat_region complement(24365..24651)
/rpt_family="Aluy"
25378..25416
/rpt_family="MER5A"
repeat_region complement(25417..25595)
/rpt_family="MER5A"
25605..25773
/rpt_family="MER5B"
25838..25906
/rpt_family="MIR"
26102..26407
/rpt_family="AluB"
repeat_region complement(26806..27100)
/rpt_family="Alusq"
27168..27438
/standard_name="174673"
repeat_region /rpt_family="(TA)n"
27565..27653
repeat_region /rpt_family="Alusq"
27660..27751
/rpt_family="L1MA9"

Query Match 92.18; Score 1093.8; DB 9; Length 173270;
Best Local Similarity 99.88; Pred. No. 3.5e-258;
Matches 1095; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 91 AGGTAACCAAGAAATGCAAGCGTCGACATCTCAGCTGCGCTGGGAACACAG 150
Db 127570 AGGTAACCAAGAAATGCAAGCGTCGACATCTCAGCTGCGCTGGGAACACAG 127629

QY 151 TCTGTGACACGAGACTCAAAATCACCCAGTCTCTCCACTGCTCTACACTGCTCT 210
Db 127630 TCTGTGACACGAGACTCAAAATCACCCAGTCTCTCTCTCCACTGCTCTACACTGCTCT 127689

QY 211 GTTTTGTGAGACTTATCACAATGGCGTGGCGATGAGATTTTCTTCAATCCGGAG 270
Db 127690 GTTTTGTGAGACTTATCACAATGGCGTGGCGATGAGATTTTCTTCAATCCGGAG 127749

QY 271 TAAATCAACTTTATTTTCTTAGACACAGTCTCTCTGATCTCTCATGATCTCT 330
Db 127750 TAAATCAACTTTATTTTCTTAGACACAGTCTCTCTGATCTCTCATGATCTCT 127809

QY 331 GACTTTTCATTCAAATTTCTAGTGATGCCAACTGGGAACAGGACACTGAGAACTTT 390
Db 127810 GACTTTTCATTCAAATTTCTAGTGATGCCAACTGGGAACAGGACACTGAGAACTTT 127869
```

```
QY 391 TGTGTGTCAGTACCTCCGTCATATTTTATTTCACAATGTATATCAGTATTTCAATCT 450
Db 127870 TGTGTGTCAGTACCTCCGTCATATTTTATTTCACAATGTATATCAGTATTTCAATCT 127929

QY 451 GGGACTGATACATATGATGCTGCTACCCAGAACACACCGGCACTTTAAACATCCACC 510
Db 127930 GGGACTGATACATATGATGCTGCTACCCAGAACACACCGGCACTTTAAACATCCACC 127989

QY 511 CAAAATCTCTGGGGGCTAGATTTCTCTGTTGTGATCTGGGCAATCTCTTACT 570
Db 127990 CAAAATCTCTGGGGGCTAGATTTCTCTGTTGTGATCTGGGCAATCTCTTACT 128049

QY 571 CTCTTGGCTTAACATGATTTGACCAACAGGACCGGAGAGACAGAAATGTGAAGAAATG 630
Db 128050 CTCTTGGCTTAACATGATTTGACCAACAGGACCGGAGAGACAGAAATGTGAAGAAATG 128109

QY 631 CTCTTTCCTTAATCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 690
Db 128110 CTCTTTCCTTAATCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 128169

QY 691 AGTCAATTTCTGATTAATTTCTTAATTTGTTATTGTATGTTATFACACTCATTAACAAAAG 750
Db 128170 AGTCAATTTCTGATTAATTTCTTAATTTGTTATTGTATGTTATFACACTCATTAACAAAAG 128229

QY 751 ACTGTACCGGTACATACGTAAGACAGAGGGGTGTAGTAAAGTCCCAAGAAAAGTGAA 810
Db 128230 ACTGTACCGGTACATACGTAAGACAGAGGGGTGTAGTAAAGTCCCAAGAAAAGTGAA 128289

QY 811 CGTCAAGTTTTCATTAATCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 870
Db 128290 CGTCAAGTTTTCATTAATCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 128349

QY 871 CCGAATTCCTTACACCCCTGACGCAAAACCGGGATGCTGTTGACTGCACTGCAAAATAC 930
Db 128350 CCGAATTCCTTACACCCCTGACGCAAAACCGGGATGCTGTTGACTGCACTGCAAAATAC 128409

QY 931 TCTGTCTATGTGAAGAGACACTCTGTGTTAACTCTTAAATGATCCCTGATCC 990
Db 128410 TCTGTCTATGTGAAGAGACACTCTGTGTTAACTCTTAAATGATCCCTGATCC 128469

QY 991 GTTCATCTATTTTCTCTGCAAGTCTCTGCAAAATCTGATTAATGATGCAAGTGT 1050
Db 128470 GTTCATCTATTTTCTCTGCAAGTCTCTGCAAAATCTGATTAATGATGCAAGTGT 128529

QY 1051 CCCCCAATCTGCAACATCTCTGCTCCAGACATAGAAAAAGAAACAGATGGTGTA 1110
Db 128530 CCCCCAATCTGCAACATCTCTGCTCCAGACATAGAAAAAGAAACAGATGGTGTA 128589

QY 1111 CCCAATGACAGACTCCATGTAACAAATTAATAAGAAATTTCAATCTCTTTGT 1170
Db 128590 CCCAATGACAGACTCCATGTAACAAATTAATAAGAAATTTCAATCTCTTTGT 128649

QY 1171 GTTCAGAACCTGTTAAA 1187
Db 128650 GTTCAGAACCTGTTAAA 128666

RESULT 10
AX369349
LOCUS AX369349 1029 bp DNA linear PAT 16-FEB-2002
DEFINITION Sequence 1 from Patent WO0202599.
ACCESSION AX369349
VERSION AX369349.1 GI:18857274
KEYWORDS
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1
Wattler,F., Wattler,S., Frommler,P. and Nehls,M.C.
Human g protein-coupled receptor igpcr17, and uses thereof
Patent: WO 0202599-A 1 10-JAN-2002;
Ingenium Pharmaceuticals AG (DE)
```

FEATURES	source	Location/Qualifiers
		1..1029
		/organism="Homo sapiens"
		/db_xref="taxon:9606"
CDS		1..1029
		/note="unamed protein product"
		/codon_start=1
		/protein_id="CAD3844.1"
		/db_xref="GI:18857275"
		/translation="MQAVNDLTSAPGNTSLCTRDYKIMQVLEPPLLTYVLEFVGLITNG LAMIYFOIRKSNFIIPIFKNTVISDLMLITPEPKLISDAKLGTCGLRFVFCOVTSV LEFTHYISIEFLGILITDRQKTRTPEKSNLNLGAKLSIVYNAFMLSIPNMA ILTRQDRDKKVKKCSFLKSEGLVMEIYNYICQVLEFIMELIYICVYLITKEIYR SYVTRGVKVPKRRKVKVKTIIIVAFVFCVPPHRAIRPIITLSQTDVDECTENIL FYVESTLMLTSLNACLDPIFYFLFCISFRNSILSMKCPNSATSIQDNKRKEBDGG DPNBERPM"
BASE COUNT	285 a	234 c 186 g 324 t
ORIGIN		
Query Match	86.6%;	Score 1027.4; DB 6; Length 1029;
Best Local Similarity	99.9%;	Pred. No. 7.8e-242;
Matches 1028; Conservative	0;	Mismatches 1; Indels 0; Gaps 0;
Oy	107	ATGCAAGCCGTGACAAATCTACCTCTCGGCTGGGAACACAGCTGTGACCGAGAAC 166
Db	1	ATGCAAGCCGTGACAAACCTCACCTCTCGGCTGGGAACACAGCTGTGACCGAGAAC 60
Oy	167	TACAAATACCCAGCAGTCTCTTCCACAGTGCCTACACGTGTCCTGTTTGTGGACTT 226
Db	61	TACAAATACCCAGCAGTCTCTTCCACAGTGCCTACACGTGTCCTGTTTGTGGACTT 120
Oy	227	ATCACAATGGCTTGGCGATGAGATTTTCTTCAATCCGAGTAATCAAACTTTAT 286
Db	121	ATCACAATGGCTTGGCGATGAGATTTTCTTCAATCCGAGTAATCAAACTTTAT 180
Oy	287	ATTTCTTAAGAACACAGTCATTTCTGATCTGTCATGATTCGATCTTCATTTCCATTC 346
Db	181	ATTTCTTAAGAACACAGTCATTTCTGATCTGTCATGATTCGATCTTCATTTCCATTC 240
Oy	347	ATTTCTTAAGATGCAAACTGGGAACAGAGACACAGAACTTTGTGTGCAATTTAC 406
Db	241	ATTTCTTAAGATGCAAACTGGGAACAGAGACACAGAACTTTGTGTGCAATTTAC 300
Oy	407	TCCGTCATATTTTATTTACAAATGTATATACGATATTCATTCCTGGGACGTAACTATC 466
Db	301	TCCGTCATATTTTATTTACAAATGTATATCAATGTATATTCATTCCTGGGACGTAACTATC 360
Oy	467	GATGCTACCGAAGAGACACAGCGCATTTAAACATCCAAACCCAAAATCTCTGGGG 526
Db	361	GATGCTACCGAAGAGACACAGCGCATTTAAACATCCAAACCCAAAATCTCTGGGG 420
Oy	527	GCTAAGATTTCTCTGTTGTGTCATCTGGGCAATTCATGTTCTTACTCTCTTGGCTAACATG 586
Db	421	GCTAAGATTTCTCTGTTGTGTCATCTGGGCAATTCATGTTCTTACTCTCTTGGCTAACATG 480
Oy	587	ATTTCGACCAACAGGACCGCGAGAGACAAAGATGTGAAGAAATGCTCTTCCCTTAATCA 646
Db	481	ATTTCGACCAACAGGACCGCGAGAGACAAAGATGTGAAGAAATGCTCTTCCCTTAATCA 540
Oy	647	GAGTCCGCTTAGTGTGGCATGAATTAATTAATTCATCTGTCAAGTCAATTTTCGGATT 706
Db	541	GAGTCCGCTTAGTGTGGCATGAATTAATTAATTCATCTGTCAAGTCAATTTTCGGATT 600
Oy	707	AATTTCTTAATTTGTATTTGTATGTATACACATTCATTAACAAAAGAACTGTACCGGCAATC 766
Db	601	AATTTCTTAATTTGTATTTGTATGTATATACACTCATTAACAAAAGAACTGTACCGGCAATC 660
Oy	767	GTAAGACGAGGGGTGTAGGTAAAGTCCCAAGAAAAGGTGAAGCTCAAAATTTCATT 826
Db	661	GTAAGACGAGGGGTGTAGGTAAAGTCCCAAGAAAAGGTGAAGCTCAAAATTTCATT 720
Oy	827	ATCATTCGTATATTTTATTTGTTTGTCTTCCATTTTGCCGAATTCCTTACCC 886

Db	721	ATCATGCTGTAATTCCTTAATTTGTTTGTGTTCCTTCCATATTTGGCCGAATTCCTACACC	780
QY	887	CTGAGCCAAACCCGGGAGTGTCTTTGACGACACTGCTGAAATATCTGTCTTGTGTGAAA	946
Db	781	CTGAGCCAAACCCGGGAGTGTCTTTGACGACACTGCTGAAATATCTGTCTTGTGTGAAA	840
QY	947	GAGAGCACTCTGTGGTAACTCTCTTAATGATCCCGGATCCGTTATCATTTTTC	1006
Db	841	GAGAGCACTCTGTGGTAACTCTCTTAATGATCCCGGATCCGTTATCATTTTTC	900
QY	1007	CTTTGCAAGTCTTCAGAAATTCCTTGAATAGATGCTGAAGTCCCAATTCGCAACA	1066
Db	901	CTTTGCAAGTCTTCAGAAATTCCTTGAATAGATGCTGAAGTCCCAATTCGCAACA	960
QY	1067	TCTGTGCTCCGAGCAATAGGAAAAAAGAACAGATGTGTGACCCCAATCAAGACT	1126
Db	961	TCTGTGCTCCGAGCAATAGGAAAAAAGAACAGATGTGTGACCCCAATCAAGACT	1020
QY	1127	CCAAATGTAA	1135
Db	1021	CCAAATGTAA	1029
RESULT	11		
LOCUS	AB052684	1029 bp	mRNA linear PRI 01-SEP-2001
DEFINITION	Homo sapiens mRNA for G1-coupled ADP receptor HOKR3, complete cds.		
ACCESSION	AB052684		
VERSION	AB052684.1	GI:14422409	
KEYWORDS			
SOURCE	Homo sapiens cDNA to mRNA.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 Takasaki,J., Kamohara,M., Saito,T., Matsumoto,M., Matsumoto,S., Ohshiki,T., Soga,T., Matsushime,H. and Furutchi,K.		
TITLE	Molecular cloning of the platelet P2Y(AC) ADP receptor: pharmacological comparison with another ADP receptor, the P2Y(1) receptor		
JOURNAL	Mol. Pharmacol.	60 (3), 432-439 (2001)	
MEDLINE	21394281		
REFERENCE	2 (bases 1 to 1029)		
AUTHORS	Takasaki,J., Kamohara,M., Saito,T., Matsumoto,M., Matsumoto,S., Ohshiki,T., Soga,T., Matsushime,H. and Furutchi,K.		
TITLE	Direct Submission		
JOURNAL	Submitted (13-DEC-2000) Jun Takasaki, Yamanouchi pharmaceutical Co., Ltd., Molecular Medicine Laboratories; 21 Miyukigoka, Tsukuba, Ibaraki 305-8585, Japan (E-mail: takasaki@yamanouchi.co.jp, Tel:81-298-52-5111, Fax:81-298-52-5444)		
FEATURES	Location/Qualifiers		
source	1..1029		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="3"		
	/map="3q21-q25"		
	1..1029		
	/gene="hork3"		
	1..1029		
	/gene="hork3"		
	/codon_start=1		
	/product="G1-coupled ADP receptor HOKR3"		
	/protein_id="BA60824.1"		
	/db_xref="GI:14422410"		
	/translation="MQAVNDNTSAPGNTSLCTRDYKITYOVLPELLYTVLFEVGLITNGI LAMIFQIRKSNFIIFLKNTVISDLMLITPEFKLISDAKLTGDLRFVCVTSV IIFVETATISISFLGIDIRQKTTREPFTSNPILGAKLSVIVAMFELSIPNN IILNRQDRDNKVKCSGLKSEFGLVMEIYNYICQVTFMNLIVICYLITELR SVYTRTEGVKPRKKVAVKVIITAVPEFCVPEHAPARIPYTLISQTRDVPDCAENTL FVYKESLIMLTSLNACLDPIITFLFCKSFNLSLMLKCPNSATSLSDNKKRKEQDGG DPMEETVW"		
gene			
CDS			
BASE COUNT	285 a	234 c	186 g 324 t
ORIGIN			

Query Match 86.6%; Score 1027.4; DB 9; Length 1029;
 Best Local Similarity 99.9%; Pred. No. 7, 8e-242;
 Matches 1028; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 107 ATGCACCCGTCGACAAATCTCACCCTGCGCTGGGAACACAGCTGTGTGACACGAGAC 166
    |||||||
DB 1 ATGCACCCGTCGACAACTCACCCTGCGCTGGGAACACAGCTGTGTGACACGAGAC 60
QY 167 TACAAATCAACCCAGGTCCTTCCCACTGCTACACTGTCCTGTTTGTGTGGACTT 226
    |||||||
DB 61 TACAAATCAACCCAGGTCCTTCCCACTGCTACACTGTCCTGTTTGTGTGGACTT 120
QY 227 ATCAAAATGGCCCTGGGATGAGGATTTCTTCAATCCGAGTAATCAACTTATT 286
    |||||||
DB 121 ATCAAAATGGCCCTGGGATGAGGATTTCTTCAATCCGAGTAATCAACTTATT 180
QY 287 ATTTTCTTAAGAACACAGTCATTTCTGATCTTCTCATGATTTGACTTTTCCATTCAA 346
    |||||||
DB 181 ATTTTCTTAAGAACACAGTCATTTCTGATCTTCTCATGATTTGACTTTTCCATTCAA 240
QY 347 ATCTTAGTGTGTCACCAACTGGGAACGACGACCTGAGAACTTTGTGTGCAAGTTACC 406
    |||||||
DB 241 ATCTTAGTGTGTCACCAACTGGGACGACGACCTGAGAACTTTGTGTGCAAGTTACC 300
QY 407 TCCGTCATATTTTATTTACAAATGTAATCAGATTTTCTGAGGACTGATTAATATC 466
    |||||||
DB 301 TCCGTCATATTTTATTTACAAATGTAATCAGATTTTCTGAGGACTGATTAATATC 360
QY 467 GATGCTACCAAGAACACGACGACGACCTTAAACATCAACCCCAAAATCTCTTGGGG 526
    |||||||
DB 361 GATGCTACCAAGAACACGACGACGACCTTAAACATCAACCCCAAAATCTCTTGGGG 420
QY 527 GCTAAGATTCCTCTGTTGTCATCTGGGCAATTCATGTTTCTTCTGCTTGGCTTAACATG 586
    |||||||
DB 421 GCTAAGATTCCTCTGTTGTCATCTGGGCAATTCATGTTTCTTCTGCTTAAACATG 480
QY 587 ATTCTGACCAACAGGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 646
    |||||||
DB 481 ATTCTGACCAACAGGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 540
QY 647 GAGTTCGCTCTAGTCTGTGCAATGTAATTAATTAATTAATTAATTAATTAATTAATTA 706
    |||||||
DB 541 GAGTTCGCTCTAGTCTGTGCAATGTAATTAATTAATTAATTAATTAATTAATTAATTA 600
QY 707 AATTTCTTAATTTGTTATGTTATGTTATGTTATGTTATGTTATGTTATGTTATGTTATG 766
    |||||||
DB 601 AATTTCTTAATTTGTTATGTTATGTTATGTTATGTTATGTTATGTTATGTTATGTTATG 660
QY 767 GTAAGAACGAGGGGTGAGGTAAAGTCCGACGACGACGACGACGACGACGACGACGAC 826
    |||||||
DB 661 GTAAGAACGAGGGGTGAGGTAAAGTCCGACGACGACGACGACGACGACGACGACGAC 720
QY 827 ATCTTCTGCTATTTCTTATTTGTTTGTCTTTCATTTTCCGCAATTTCTTACACC 886
    |||||||
DB 721 ATCTTCTGCTATTTCTTATTTGTTTGTCTTTCATTTTCCGCAATTTCTTACACC 780
QY 887 CTGAGCCAAACCCGAGGATGCTTTGACTGACGCTGGAATACTGTTCTATGTGAAA 946
    |||||||
DB 781 CTGAGCCAAACCCGAGGATGCTTTGACTGACGCTGGAATACTGTTCTATGTGAAA 840
QY 947 GAGAGCACTGTGTTGTTTACCTCTTAAATGCAATGCTGATGCTGATGCTGATGCTGAT 1006
    |||||||
DB 841 GAGAGCACTGTGTTGTTTACCTCTTAAATGCAATGCTGATGCTGATGCTGATGCTGAT 900
QY 1007 CTTTGCAAGTCCTTCACAAATTCCTTGAATGCTGTAAGTGTGTAAGTGTGTAAGTGTGTA 1066
    |||||||
DB 901 CTTTGCAAGTCCTTCACAAATTCCTTGAATGCTGTAAGTGTGTAAGTGTGTAAGTGTGTA 960
QY 1067 TCTCTGTCCAGACATATAGGAAAAAAGAACAGAGATGCTGTGACCCCAATGGAAGAGAT 1126
    |||||||
DB 961 TCTCTGTCCAGACATATAGGAAAAAAGAACAGAGATGCTGTGACCCCAATGGAAGAGAT 1020

```

QY 1127 CCAATGTAA 1135
 |||||||
 DB 1021 CCAATGTAA 1029

RESULT 12
 AB083596 1029 bp DNA linear PRI 24-MAY-2002
 LOCUS
 DEFINITION Homo sapiens GPCR gene for putative G-protein coupled receptor,
 complete CDS, clone:hgPCR14.
 ACCESSION AB083596
 VERSION AB083596.1 GI:20152255
 KEYWORDS
 SOURCE Homo sapiens
 ORGANISM Homo sapiens DNA, clone:hgPCR14.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS Takeda,S., Kadowaki,S., Haga,T., Takaesu,H. and Mitaku,S.
 TITLE Identification of G protein-coupled receptor genes from the human genome sequence
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1029)
 AUTHORS Takeda,S., Kadowaki,S., Haga,T., Takaesu,H. and Mitaku,S.
 TITLE Direct Submission
 JOURNAL Submitted (10-APR-2002) Shigeaki Takeda, Gunma University,
 Department of Biological and Chemical Engineering, Faculty of
 Engineering, 1-5-1, Kiryu, Gunma 376-8515, Japan
 (E-mail:stakeda@bce.gunma-u.ac.jp, Tel:+81-277-30-1434,
 Fax:+81-277-30-1434)

FEATURES
 source
 1..1029
 location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="hgPCR14"
 1..1029
 /gene="GPCR"
 1..1029
 /gene="GPCR"
 /note="predicted with SOSI analysis"

CDS
 /product="putative G-protein coupled receptor"
 /protein_id="BAB89309.1"
 /db_xref="GI:20152255"
 /translation="MOAVDLISAPGNTSLCTPDYKITQVLPILATYVLFVGLNLG
 LAMRIFQISKSNFIFLKNYISDLMLITRPYILISAKIGLPLRFVQVTSV
 IFPTMISIFLGLTIDRYQRTTPKTSNKNLGLAKISLVITAEFELISLPM
 ILTNRPDRKNVSKSEFLKSEFLVHEIVNYICQYIFWLNFLIVCYLLITRELYR
 SYVRGQGVKVPKRNKVFIIIAVEFLFVPHFARILPYLSQTRDVEDCAENYL
 FYVEESTLMTISLNACIDPRIFYFLKSFNSLSLSMLKCPNSATLSLSQDRKKEQDG
 DPHEETPM"

BASE COUNT 285 a 234 c 185 g 325 t
 ORIGIN

Query Match 86.4%; Score 1025.8; DB 9; Length 1029;
 Best Local Similarity 99.8%; Pred. No. 1.9e-241;
 Matches 1027; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 107 ATGCACCCGTCGACAAATCTCACCCTGCGCTGGGAACACAGCTGTGTGACACGAGAC 166
    |||||||
DB 1 ATGCACCCGTCGACAACTCACCCTGCGCTGGGAACACAGCTGTGTGACACGAGAC 60
QY 167 TACAAATCAACCCAGGTCCTTCCCACTGCTACACTGTCCTGTTTGTGTGGACTT 226
    |||||||
DB 61 TACAAATCAACCCAGGTCCTTCCCACTGCTACACTGTCCTGTTTGTGTGGACTT 120
QY 227 ATCAAAATGGCCCTGGGATGAGGATTTCTTCAATCCGAGTAATCAACTTATT 286
    |||||||
DB 121 ATCAAAATGGCCCTGGGATGAGGATTTCTTCAATCCGAGTAATCAACTTATT 180
QY 287 ATTTTCTTAAGAACACAGTCATTTCTGATCTTCTCATGATTTGACTTTTCCATTCAA 346
    |||||||
DB 181 ATTTTCTTAAGAACACAGTCATTTCTGATCTTCTCATGATTTGACTTTTCCATTCAA 240

```


Db	482	GTATGCTAAACTGGAGACCGGGCCCTGTGAGAACCTTGCTGGCCAACTGTACTTCCATGCA	541
Qy	414	TATTTTATTTTCCAAATGTATATTCAGTATTTTCATTCCTCTGGAGCTGATATCTATCGATGCT	473
Db	542	CAATTTTATTTTACAAATGATATATCATGTATATGTTCCTGGGGGTGTGATTAACATATGACCGCT	601
Qy	474	ACCAAGAACCCACACAGGCGCATTTTAAAAATCTCCAAACCCCAAAATATCTCTTGGGGGCTAGA	533
Db	602	ACCTGAAGACCAACACAGGCGCATTTTAAAAAGTCCAGGCCAGCAATCTCTTGGGTGCANAAG	661
Qy	534	TTTCTCTCTGTTGTGATCTCTGGGCATTCATGTTCTTACTCTCTTGGCTAACATGATTTCTGA	593
Db	662	TTCTCTCTGTTGTGATCTCTGGGCATTCATGTTCTTACTCTCTTGGCTAACATGATTTCTGA	721
Qy	594	CCAAACAGCACCACGAGACACAAGATGTGAAGAAATGCTCTTCTCTTAAACAGAGTTG	653
Db	722	CCAAACAGCACCACGAGACACAAGATGTGAAGAAATGCTCTTCTCTTAAACAGAGTTG	781
Qy	654	GTCATGTCGACATGAATAGTAAATTTACATCTGTCAAGTCATTTTCTGATTTATTTCT	713
Db	782	GTCATGTCGACATGAATAGTAAATTTACATCTGTCAAGTCATTTTCTGATTTATTTT	841
Qy	714	TAAATTTGTTATGTATGTATATACACTCATTTACAAAAAGAACTGTACCGGTCATACGTAGAA	773
Db	842	TAAATTTGTTATGTATGTATATACACTCATTTACAAAAAGAACTGTACCGGTCATACGTAGAA	901
Qy	774	CGAGGGGTGTAGGTAAATGCCACAGAAAAAGGTGACAGTCMAAGTTTTCATTTATCATTG	833
Db	902	CGAGGGGTGTAGGTAAATGCCACAGAAAAAGGTGACAGTCMAAGTTTTCATTTATCATTG	961
Qy	834	CTGTATTTCTTATTTTGTGTTTGTCTTTCATTTTGCATTTTGGCCGAATTCCTTACACCCGTGACC	893
Db	962	CTGTATTTCTTATTTTGTGTTTGTCTTTCATTTTGCATTTTGGCCGAATTCCTTACACCCGTGACC	1021
Qy	894	AAACCCGGGAGTCTCTTGTACGTGACCTGCTGAAAAATACCTCTGTGTATGTGAAGAGACGA	953
Db	1022	AAACCCGGGAGTCTCTTGTACGTGACCTGCTGAAAAATACCTCTGTGTATGTGAAGAGACGA	1081
Qy	954	CTCTGTGGTTAACTTCTTAAATGCAATGCCCTGTGATCCGTTTCATCTATTTTTCCTTTTCA	1013
Db	1082	CCCTATGTGCTGACGTCATGTGAACGCCCTGCTGTATCATTCATCTATTTTTCCTTTTCA	1141
Qy	1014	AGTCTTTCAGAAATTCCTTGTATAGTATGCTGGAAGTGGCCCAATTCCTGCAACATCTCTGT	1073
Db	1142	AGTCTTTCAGAAATTCCTTGTACAAAGCATGCTGAGGTGCTCAAAACCT--ACATTCACAT	1198
Qy	1074	CCCAAGGACAAATGCAAAAAAACAAGATGTGTGATGCCCAAAATGAAGAGACTCCAAATGT	1133
Db	1199	CTGGGACAAACAAGAGAGAAAGACAAAGAGTGGCGAACAACGCAAGACAGACCCCAATGT	1258
Qy	1134	AAACAAATTAACATGAAGAAATATTTTCAATCTCT	1166
Db	1259	AGA-ACATTTACCCAAAGGGGCTGCTCACTCTTT	1290
RESULT 14			
BC025428			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 Contact: MGC help desk
 Email: egads-remail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 CNA Library Preparation: Life Technologies, Inc.
 CNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
 Gaithersburg, Maryland,
 Web Site: <http://www.nisc.nih.gov/>
 Contact: nisc.mgc@nih.gov
 After, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Grante, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kong, P., Latic, P., Legaspi, R.,
 Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stattin, P., Thomas, P.J., Touchman, J.W.,
 Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggin, L.,
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAC Plate: 39 Row: 1 Column: 4
 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

FEATURES
source
location/Qualifiers
1..2422
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="C2BCH 11"
/clone="MGC:29342 IMAGE:5035481"
/tissue_type="Mammary tumor metastasized to lung. Tumor arose spontaneously from a senescent normal mammary (clonal) outgrowth infected with the virus MMTV."
/clone_id="NCI-GAP.Lu29"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"
422..1465
/codon_start=1
/product="unknown (protein for MGC:29342)"
/protein_id="AAH25428.1"
/db_xref="GI:19343841"
/translation="MDMPGVNTTSANTFSPGSTLQVDKYITQVLFPLLYVLEFA
GLTNSVLAATFQIRKSNFIEFLKNTVSDLMITLFFPKLISDAKLGACPLRTLY
CQVTSVFYFTMYTISIFGLITIDRLTKTRPKFTSSPSNLGAKLSVIVAMFL
ISLPMNLTRRPKRDQVTCSPFLSGFLVWHEIVYICQVTFMNFELIVCYSLI
TKELRSYVTRGSAKPVKKVKKVYKVFIIAVERICVCPHAPRIPTLSQTAVERC
SAERTLPEYKESTIMLTSLNACLDPIFYFLCKSFNLSLSMLCSNSTISGTNKK
GOREGESEETPM"

BASE COUNT 781 a 523 c 419 g 699 t
ORIGIN

Query Match 61.7% Score 732.2 DB 10 Length 2422;
 Query Local Similarity 82.2% Pred. No. 2,76-169;
 Matches 866; Conservative 0; Mismatches 183; Indels 4; Gaps 2;

114 CCGTCGACAAATCTCACTCTGCGCTGGGAACACCACTCTGTGCACGAGACTACAA 173
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 447 CCTGAGCAATACCACTCTTCCCTGGGACGAGACCCTGTGGGTGAGACATACAGA 506
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 174 TCACCCAGAGTCCTTCCCACTGCTCTACACTGTGTCGCTGTTTGTGTGGACTTATAC 233
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 507 TCACCCAGAGTCCTTCCCACTGCTCTACACCGTCGCTGTTTGTGGGTGCTCATACAGA 566
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 234 ATGGCTGGCGATAGAGATTTTCTTCAAACTCCGAGATTAATCAAACTTATATTTTTC 293
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 567 ACAGCTGGCAATGAGGATTTTCTTTCGATTCGCGAGTAATTCACATCATCATTTTTC 626
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 294 TTAAAGACACAGTATTTTGATCTTCTCATGATTTGACTTTTCCATCAAAATTCCTTA 353
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 627 TTAAAGACACAGTATTTTGATCTTCTCATGATTTGACTTTTCCATCAAAATTCCTTA 686

[illegible]

```

JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 211964)
AUTHORS      McPherson,J.D. and Waterston,R.H.
TITLE        Direct Submission
JOURNAL      Submitted (21-MAY-2002) Genome Sequencing Center, 4444 Forest Park
              Parkway, St. Louis, MO 63108, USA
              3 (bases 1 to 211964)
REFERENCE    McPherson,J.D. and Waterston,R.H.
AUTHORS      Direct Submission
TITLE        Submitted (27-MAY-2002) Genome Sequencing Center, 4444 Forest Park
JOURNAL      Parkway, St. Louis, MO 63108, USA

COMMENT      ----- Genome Center -----
              Center: Washington University Genome Sequencing Center
              Center code: WUGSC
              Web site: http://genome.wustl.edu/gsc/index.shtml
              Contact: submissions@wustl.wustl.edu
              ----- Project Information -----
              Center project name: M_BB0427G24
              ----- Summary Statistics -----
              Sequencing vector: Plasmid; 0%
              Chemistry: Dye-primer EM; 0% of reads
              Assembly program: Dye-terminator Big Dye; 100% of reads
              Consensus quality: Phrap; version 0.990319
              Consensus quality: 210721 bases at least Q40
              Consensus quality: 211170 bases at least Q30
              Consensus quality: 211464 bases at least Q20
              Insert size: 195000; agarose-fp
              Insert size: 211764; sum-of-contigs
              Quality coverage: 12.85 in Q20 bases; sum-of-contigs
              Quality coverage: 10.63 in Q20 bases; sum-of-contigs
              -----
              * NOTE: This is a 'working draft' sequence. It currently
              * consists of 3 contigs. The true order of the pieces
              * is not known and their order in this sequence record is
              * arbitrary. Gaps between the contigs are represented as
              * runs of N, but the exact sizes of the gaps are unknown.
              * This record will be updated with the finished sequence
              * as soon as it is available and the accession number will
              * be preserved.
              *
              1      1402: contig of 1402 bp in length
              *      1403      1502: gap of unknown length
              *      1503      91430: contig of 89928 bp in length
              *      91431      91530: gap of unknown length
              *      91531      211964: contig of 120434 bp in length.
              *
              Location/Qualifiers
                1..211964
                  /organism="Mus musculus"
                  /db_xref="taxon:10090"
                  /chromosome="UNK"
                  /clone="RP24-427G24"
                  1..1402
                    /note="assembly_name:Contig5"
                    1503..91430
                      /note="assembly_name:Contig6"
                      91531..211964
                        /note="assembly_name:Contig7"
BASE COUNT   62887 a 43653 c 43160 g 62052 t      212 others
ORIGIN
Query Match      61.7%; Score 732.2; DB 2; Length 211964;
Best Local Similarity 82.2%; Pred. No. 2.5e-16;
Matches 866; Conservative 0; Mismatches 183; Indels 4; Gaps 2;

QY      114 CCGTGCACAATTCACCTCTCGGCGCTGGGAACACAGTCTGTGCACACAGAGACTCAAAA 173
          || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
DB 108731 CCTACGCCAATTCACACCTCTTCCTCGGACACAGACCCCTGTGCGCTACAGACTCAACA 108790

QY      174 TCACCCAGAGTCTCTCCACATGCTCTACACTGTGCTCTTTTGTGGACTTATACAA 233
          |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
DB 108791 TCACCCAGAGTCTCTCCACATGCTCTACACTGTGCTCTTTTGTGGACTTATACAA 108850

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2002, 22:50:25 ; Search time 322 Seconds
(without alignments)
8301.621 Million cell updates/sec

Title: US-09-835-922-1
1187
1 ggtaccgcgggacccctctag.....tggttcagaactcgttaaa 1187

Sequence: IDENTITY NUC

Scoring table: Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*
1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1187	100.0	1187	24	ABAG2641 Human SP168 receptor
2	1160	97.7	1893	22	AAK52382 Human polynucleoti
3	1156.8	97.5	2249	20	AAV69760 EBV-induced G-prot
4	1156.8	97.5	2286	21	AAAG7126 Human 15625 recept
5	1153.4	97.2	1332	24	AAAG8079 Human DNA for pote
6	1151.4	97.0	1330	22	AAAD08695 Human P2-purifierg1
7	1137.4	95.8	1328	22	AAAD08700 Human P2-purifierg1
8	1108.8	93.4	2272	21	AAA27127 Macaque ortholog o
9	1093.8	92.1	1101	24	AA168803 Human PAFR-3 DNA f

10	1093.8	92.1	1143	24	ABK50286 Human DNA encoding
11	1093.8	92.1	2245	24	AA168802 Human PAFR3 DNA.
12	1029	86.7	1029	21	AAAD01133 Human orphan G pro
13	1029	86.7	1029	21	AAA46034 Human G protein co
14	1027.4	86.6	1029	24	AB159205 Nucleotide sequenc
15	1027.4	86.6	1029	24	ABA85535 Human G protein-co
16	1022.4	86.1	1107	22	AAAD08694 Human P2-purifierg1
17	729.8	61.5	1044	24	ABA85535 Murine G protein-c
18	701.4	59.1	1289	22	AAAD08693 Rat P2-purifiergic
19	686.2	57.8	2711	21	AAA48642 Rat MP-10 receptor
20	511	43.0	699	22	ABA09444 Human EBV-induced
21	511	43.0	699	22	AAK53366 Human polynucleoti
22	466	39.3	466	24	ABA95338 Partial human G pr
23	428.8	36.1	433	20	AAK39462 Human secreted pro
24	407.6	34.3	542	24	ABA94314 Platelet activatin
25	407.6	34.3	542	24	AAK39462 Human DNA for pote
26	307.2	25.9	1002	22	AAK39462 Novel human G prot
27	307.2	25.9	1002	22	AAK39462 Human G protein co
28	307.2	25.9	1484	22	AAAD06503 Human CON203 G pro
29	307.2	25.9	1812	21	AAK39462 Human chemokine re
30	307.2	25.9	2760	24	AAK39462 Human DNA for pote
31	305.6	25.7	1002	21	AAAD01126 Human orphan G pro
32	305.6	25.7	1124	20	AAK35385 Human G protein co
33	305.6	25.7	1124	20	AAK35385 CDNA clone HNEA81
34	305.6	25.7	1124	22	AAK35385 Nucleotide sequenc
35	305.6	25.7	1124	23	AAK35385 Human HNEA81 CDNA
36	304	25.6	1488	19	AAV59144 Human chemokine re
37	302.4	25.5	1488	24	AA172599 Chemokine receptor
38	295.4	24.9	1474	17	AA173390 Human G-protein co
39	295.4	24.9	1474	21	AAV0770 Human G-protein co
40	295.4	24.9	1474	21	AAZ60118 Human G-protein co
41	294.8	24.8	1591	24	ABL90619 Human polynucleoti
42	279	23.5	302	20	AAK51843 Human secreted pro
43	263.2	22.2	356	24	ABA98542 Partial murine G p
44	253	21.3	1011	22	AAK57695 Rat 7-transmembran
45	252.4	21.3	2416	24	ABN86182 Human stem cell G-

ALIGNMENTS

RESULT 1	
ABAG2641	
ID	ABAG2641 standard; CDNA: 1187 BP.
AB	
XX	
XX	ABAG2641;
AC	
XX	
XX	25-MAR-2002 (first entry)
DT	
XX	
DE	Human SP168 receptor encoding CDNA SEQ ID NO:1.
XX	
KW	Human: SP168 receptor; mammalian G-protein coupled receptor; GPCR;
KW	neurodegenerative disorder; Parkinson's disease; Alzheimer's disease;
KW	Huntington's disease; amyotrophic lateral sclerosis; ALS; MS; receptor;
KW	multiple sclerosis; gene; ss.
KW	
XX	
OS	Homo sapiens.
XX	
XX	
FT	Key
FT	location/qualifiers
FT	107..1135
FT	/*tag= a
FT	/product= "Sp168 receptor"
XX	
XX	
XX	US2001046497-A1.
PD	
XX	29-NOV-2001.
XX	
XX	16-APR-2001; 2001US-0835922.
PF	
XX	
XX	21-APR-2000; 2000US-199044P.
PR	
XX	
XX	(ZHANG) ZHANG F L.
PA	(LUDOL) LUDOL L.

PA (GUST/) GUSTAFSON E.
 PA (LIU/) LIU Y.
 PA (CHEN/) CHEN G.
 PI Zhang FL, Luo L, Gustafson E, Liu Y, Chen G;
 DR WPI: 2002-082414/11.
 DR P-PSDB; AB805031.
 XX
 PT Identifying modulators of mammalian G-Protein Coupled Receptor SP168,
 PT useful for treating Parkinson's Disease, Alzheimer's Disease,
 PT Huntington's Disease, amyotrophic lateral sclerosis and multiple
 PT sclerosis -
 XX
 PS Example 1; Page 13-14; 16pp; English.

CC The present invention describes a method for identifying agonists and
 CC antagonists of mammalian G-Protein Coupled Receptor (GPCR) SP168 which
 CC may be used to treat neurodegenerative disorders. The method comprises:
 CC (a) contacting a mammalian SP168 receptor (or a functional fragment) in
 CC the presence of a known amount of a labeled SP168 receptor ligand with
 CC a sample to be tested for the presence of the SP168 receptor ligand and
 CC antagonist; and (b) measuring the amount of labeled SP168 ligand
 CC specifically bound to the receptor (the SP168 receptor agonist or
 CC antagonist in the sample is identified by measuring the difference in
 CC binding of the labeled SP168 receptor ligand to the receptor, compared
 CC to what would be measured in the absence of such agonist or antagonist).
 CC The method is used to detect agonists and antagonists (especially
 CC antibodies) of the SP168 GPCR which may be used to treat a
 CC neurodegenerative disorder, such as Parkinson's Disease, Alzheimer's
 CC Disease, Huntington's Disease, amyotrophic lateral sclerosis (ALS) and
 CC multiple sclerosis (MS) in mammals, especially humans. The present
 CC sequence encodes the human SP168 receptor which can be used in the
 CC method of the invention.

SQ Sequence 1187 BP; 340 A; 270 C; 212 G; 365 T; 0 other;

Query Match 100.0%; Score 1187; DB 24; Length 1187;
 Best Local Similarity 100.0%; Pred. No. 1e-304;
 Matches 1187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTACCCGGGGATCTCTAGAGTCGACCTGACGCTGCAATTAATCTACTACTGATGATA 60
 DB 1 GGTACCCGGGGATCTCTAGAGTCGACCTGACGCTGCAATTAATCTACTACTGATGATA 60
 QY 61 CATTCGAACCCCTCCAGAAATGACATGATGATGATGATGATGATGATGATGATGATGATGAT 120
 DB 61 CATTCGAACCCCTCCAGAAATGACATGATGATGATGATGATGATGATGATGATGATGATGAT 120
 QY 121 CAATCTACCTCTGCGCTGGGAAACACAGTCTGTGACACAGACATGACAAATATCATCCA 180
 DB 121 CAATCTACCTCTGCGCTGGGAAACACAGTCTGTGACACAGACATGACAAATATCATCCA 180
 QY 181 GGTCTCTTCCCACTGCTACACTGTCTGTTTGTGGAAGTTATGCAAAATGGCCT 240
 DB 181 GGTCTCTTCCCACTGCTACACTGTCTGTTTGTGGAAGTTATGCAAAATGGCCT 240
 QY 241 GGGATGAGGATTTCTTCAATCCGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
 DB 241 GGGATGAGGATTTCTTCAATCCGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
 QY 301 CACAGTCAATTTCTGATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
 DB 301 CACAGTCAATTTCTGATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
 QY 361 CAAACTGGGAACAGACACATGAGAACTTTGTGTGCAAGTATCTCCCTCATATATTTTA 420
 DB 361 CAAACTGGGAACAGACACATGAGAACTTTGTGTGCAAGTATCTCCCTCATATATTTTA 420
 QY 421 TTTCACAAATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 DB 421 TTTCACAAATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480

QY 481 GACCACCAAGGCGATTTAAACATCCAAACCCCAAAATCTCTGTGGGGCTAAGATTCCTC 540
 DB 481 GACCACCAAGGCGATTTAAACATCCAAACCCCAAAATCTCTGTGGGGCTAAGATTCCTC 540
 QY 541 TGTGTGATCAGGCGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 DB 541 TGTGTGATCAGGCGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 QY 601 GCACCCGAGAGACAAAGATGTAAGAAATGCTCTTCTTAAATCAGAGTTCGGTCTAGT 660
 DB 601 GCACCCGAGAGACAAAGATGTAAGAAATGCTCTTCTTAAATCAGAGTTCGGTCTAGT 660
 QY 661 CTGGCAGTAATATGTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
 DB 661 CTGGCAGTAATATGTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
 QY 721 TATGTATGTTAACTACTATTAACAAAGAGTACCGGTCATAGTAAAGCAGAGGG 780
 DB 721 TATGTATGTTAACTACTATTAACAAAGAGTACCGGTCATAGTAAAGCAGAGGG 780
 QY 781 TGTAGTAAAGTCCCGAGAAAAGTGAACGTCAAGTTTCATTAATGCTGATAT 840
 DB 781 TGTAGTAAAGTCCCGAGAAAAGTGAACGTCAAGTTTCATTAATGCTGATAT 840
 QY 841 CTATATTTGTTTGTCTTCTTCAATTTGCGCCGAATTCCTTACACCTGAGCCAAACCCG 900
 DB 841 CTATATTTGTTTGTCTTCTTCAATTTGCGCCGAATTCCTTACACCTGAGCCAAACCCG 900
 QY 901 GGATGCTTTGACGCTGACCTGCTGAAATATCTGTTCTATGTAAGAGAGACCTGCTGTG 960
 DB 901 GGATGCTTTGACGCTGACCTGCTGAAATATCTGTTCTATGTAAGAGAGACCTGCTGTG 960
 QY 961 GTTAACCTCTTAAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 1020
 DB 961 GTTAACCTCTTAAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 1020
 QY 1021 CAGAAATTCCTGTAAGTATGCTGAAGTCCCAATTCGCAACATCTCTGCAAGTCCCT 1080
 DB 1021 CAGAAATTCCTGTAAGTATGCTGAAGTCCCAATTCGCAACATCTCTGCAAGTCCCT 1080
 QY 1081 CAATAGGAAAAAAGACAGAGATGCTGACCCCAATGAGAGACATGATGATGATGATGATGATGAT 1140
 DB 1081 CAATAGGAAAAAAGACAGAGATGCTGACCCCAATGAGAGACATGATGATGATGATGATGATGAT 1140
 QY 1141 TTAACTAAGGAAATATTTCAATCTCTTGTGTCAGAACTGCTTAA 1187
 DB 1141 TTAACTAAGGAAATATTTCAATCTCTTGTGTCAGAACTGCTTAA 1187

RESULT 2
 AAK52382
 ID AAK52382 standard; cDNA; 1893 BP.
 XX
 AC AAK52382;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polynucleotide seq. ID NO 927.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukemia;
 KW nervous system disorder; arthritis; inflammation; ss.
 OS Homo sapiens.
 XX
 PN MO200157190-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US04098.
 XX
 PR 03-FEB-2000; 2000US-0496914.

PR 27-APR-2000; 2000US-0560875.
 PR 20-JUN-2000; 2000US-0598075.
 PR 19-JUL-2000; 2000US-0620325.
 PR 01-SEP-2000; 2000US-0654936.
 PR 15-SEP-2000; 2000US-0663561.
 PR 20-OCT-2000; 2000US-0693325.
 PR 30-NOV-2000; 2000US-0728422.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YF, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 PI Zhuo QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AH, Yang Y, Wehrman T, Goodrich R;
 XX
 DR WPI: 2001-476283/51.
 DR P-PSDB; AAM79249.
 XX
 PT Nucleic acids encoding polypeptides with cytokine-like activities,
 PT useful in diagnosis and gene therapy -
 XX
 PS Claim 1; Page 3051-3052; 6221pp; English.
 XX
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoietic regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC treatment of cancer, leukemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAM80020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.
 CC
 XX
 SQ Sequence 1893 BP: 620 A; 393 C; 318 G; 560 T; 2 other:
 Query Match 97.7%; Score 1160; DB 22; Length 1893;
 Best Local Similarity 100.0%; Pred. No. 1.8e-297;
 Matches 1160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 28 CTGACGCTGCAATATCTACTACTGATATCAAAACCTCTCGAATCAACAGTT 87
 Db CTGACGCTGCAATATCTACTACTGATATCAAAACCTCTCGAATCAACAGTT 269
 QY 88 ATGAGGTACCAACAAAGAAATGCAACCGGTGACAACTCTCACTCTGGCGGAAAC 147
 Db 270 ATCAGGTACCAACAAAGAAATGCAACCGGTGACAACTCTCACTCTGGCGGAAAC 329
 QY 148 CAGTCTGTGACACAGAGACTCAAAATACCCAGGCTCTCCACAGCTCTACACTGT 207
 Db 330 CAGTCTGTGACACAGAGACTCAAAATACCCAGGCTCTCTCCACAGCTCTACACTGT 389
 QY 208 CCGTGTCTTGTGTGACTATACAAATGCGCTGGCATGAGGATTTCTTTCAATCCG 267
 Db 390 CCGTGTCTTGTGTGACTATACAAATGCGCTGGCATGAGGATTTCTTTCAATCCG 449
 QY 268 GAGTAAATCAAACTTATTTCTTAAGAACACAGCTCTTGTGATCTTCATCATAT 327
 Db 450 GAGTAAATCAAACTTATTTCTTAAGAACACAGCTCTTGTGATCTTCATCATAT 509
 QY 328 TCTGACTTTTCCATCAAAATTTCTAGTATGCCAAACTGGGAAACAGACACTGAGAAC 387
 Db 510 TCTGACTTTTCCATCAAAATTTCTAGTATGCCAAACTGGGAAACAGACACTGAGAAC 569
 QY 388 TTTTGTGTGCTCAAGTATCTCCGTCATATTTTATTTTCAATGATGATAGATTTTCAT 447
 Db 570 TTTTGTGTGCTCAAGTATCTCCGTCATATTTTATTTTCAATGATGATAGATTTTCAT 629
 QY 448 CTTGGAGCTGATTAATCTATGATCGCTACAGAGACCAAGGCGCATTTAAACATCCAA 507

Db 630 CTTGGAGCTGATTAATCTATGATCGCTACAGAGACCAACAGGCCATTTAAACATCCAA 689
 QY 508 CCCCCAAAATCTCTGGGGGGCTAGATTTCTCTGTGTGATCTGGGCAATTCATGTCCT 567
 Db 630 CCCCCAAAATCTCTGGGGGGCTAGATTTCTCTGTGTGATCTGGGCAATTCATGTCCT 749
 QY 568 ACTCTCTTGGCTTAACATGATTTCTGACCAACAGGACCGAGAGACAAAGATGTGAAGA 627
 Db 750 ACTCTCTTGGCTTAACATGATTTCTGACCAACAGGACCGAGAGACAAAGATGTGAAGA 809
 QY 628 ATGCTCTTCTCTTAATATGAGAGTGGTCTGAGTCTGCGCATGAAATGATTAATTCATCTG 687
 Db 810 ATGCTCTTCTCTTAATATGAGAGTGGTCTGAGTCTGCGCATGAAATGATTAATTCATCTG 869
 QY 688 TCAAGTCAATTTCTGATTAATTTCTTAATGTTATGTTATGATGTTATGATGATTAATCAAA 747
 Db 870 TCAAGTCAATTTCTGATTAATTTCTTAATGTTATGTTATGATGTTATGATGATTAATCAAA 929
 QY 748 AGAAGTGAACGGGTATGATGATGAAGAGAGGGGTGATGATTAAGTCCCGAGAGAAAAGT 807
 Db 930 AGAAGTGAACGGGTATGATGATGAAGAGAGGGGTGATGATTAAGTCCCGAGAGAAAAGT 989
 QY 808 GAGCTCAAAAGTTTCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 867
 Db 990 GAGCTCAAAAGTTTCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1049
 QY 868 TGCCCGAATTCCTTACACCCGAGCCAAACCGGGATGTCCTGACGACGTCGAAAA 927
 Db 1050 TGCCCGAATTCCTTACACCCGAGCCAAACCGGGATGTCCTGACGACGTCGAAAA 1109
 QY 928 TACTGTCTTAATGTAAGAAAGAGACGCTGTGTTAACTCTTAAATGATGATGATGATGAT 987
 Db 1110 TACTGTCTTAATGTAAGAAAGAGACGCTGTGTTAACTCTTAAATGATGATGATGATGAT 1169
 QY 988 TCCGTTCAATATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1047
 Db 1170 TCCGTTCAATATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1229
 QY 1048 GTGCCCAATTCGCAACATCTCTGTCCAGAGCAATGAGAAAGAAACAGAGATGTGG 1107
 Db 1230 GTGCCCAATTCGCAACATCTCTGTCCAGAGCAATGAGAAAGAAACAGAGATGTGG 1289
 QY 1108 TGACCCAAATGAGAGACTCCATGTAACAAATTAAGTAAGAAATATTTCAATCTCTT 1167
 Db 1290 TGACCCAAATGAGAGACTCCATGTAACAAATTAAGTAAGAAATATTTCAATCTCTT 1349
 QY 1168 TGTGTTCAAGACTCGTTAAA 1187
 Db 1350 TGTGTTCAAGACTCGTTAAA 1369
 RESULT 3
 AAV69760
 ID AAV69760 standard; cDNA: 2249 BP.
 XX AAV69760;
 XX 22-FEB-1999 (first entry)
 DE EBV-induced G-protein coupled receptor (EBI-2) encoding cDNA.
 XX
 XX EBV-induced G-protein coupled receptor (EBI-2) encoding cDNA.
 KW EBV-induced G-protein coupled receptor; EBI-2; Epstein-Barr Virus; ulcer;
 KW endothelium-differentiation gene; EDS-1-like G-protein coupled receptor;
 KW recombinant; agonist; asthma; Parkinson's disease; heart failure; asthma;
 KW hypotension; urinary retention; osteoporosis; antagonist; hypertension;
 KW angina pectoris; myocardial infarction; allergy; psychosis; depression;
 KW migraine; vomiting; stroke; eating disorder;migraine headache; cancer;
 KW prostatic hypertrophy; detection; drug screening; ss.
 OS Homo sapiens.
 XX
 XX
 FH Key Location/Qualifiers
 FT CDS 226..1255

```

FT      /tag= a
FT      /transl_except= (pos:241..243; aa:Asn)
FT      /transl_except= (pos:265..267; aa:Thr)
FT      /product= "EBV-induced G-protein coupled receptor"
XX      MO950549-A2.
XX      12-NOV-1998.
XX      07-MAY-1998; 98WO-0509048.
XX      07-MAY-1997; 97US-0852824.
XX      (HUMA-) HUMAN GENOME SCI INC.
XX      IL Y, Ruben SM;
XX      WPI: 1999-034722/03.
XX      P-PSDB; AAW81576.
XX      New isolated human G-protein coupled receptors - used to develop
XX      products for treating e.g. asthma, Parkinson's disease, heart
XX      failure, osteoporosis, hypertension, psychoses, eating disorders or
XX      cancers
XX      Claim 9; Fig 1A-C; 65pp; English.
XX      This cDNA encodes a EBV-induced G-protein coupled receptor (EBI-2)
XX      polypeptide. The present sequence is deposited under the accession number
XX      ATCC No: 209003. The invention provides two human G-protein coupled
XX      receptor polypeptides. The polypeptides are human Epstein-Barr Virus
XX      (EBV)-induced G-protein coupled receptor, designated EBI-2 polypeptide
XX      and a human endothelium-differentiation gene (EDG) like G-protein coupled
XX      receptor, designated EDG-1-like G-protein coupled receptor. Vectors
XX      comprising the EBI-2 and EDG-1-like polypeptides encoding DNA can be used
XX      to transform host cells for the recombinant production of the proteins.
XX      Agonists for G-protein coupled receptors can be used for the treatment of
XX      asthma, Parkinson's disease, acute heart failure, hypotension, urinary
XX      retention and osteoporosis. Antagonists can be used for the treatment of
XX      hypertension, angina pectoris, myocardial infarction, ulcers, asthma,
XX      allergies, psychoses, depression, migraine, vomiting, stroke, eating
XX      disorders, migraine headaches, cancer and benign prostatic hypertrophy.
XX      The products can also be used for detection, diagnosis and drug
XX      screening.
XX      Sequence 2249 BP; 784 A; 420 C; 347 G; 697 T; 1 other;
SQ
Query Match          97.5%; Score 1156.8; DB 20; Length 2249;
Best Local Similarity 99.8%; Pred. No. 1.4e-296;
Matches 1158; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY      28 CTGCAAGCTGCAATTAAGTACTTACTGATGATCAATTCAAACCCCTCCAGAAATCAAGATT 87
DB      147 CTGCAAGCTGCAATTAAGTACTTACTGATGATCAATTCAAACCCCTCCAGAAATCAAGATT 206
QY      88 ATGAGGTAAACCAAGAAATGCAAGCCGTCGCAATCTCACCTCTGCGCTGGGAACAC 147
DB      207 ATGAGGTAAACCAAGAAATGCAAGCCGTCGCAATCTCACCTCTGCGCTGGGAACAA 266
QY      148 CAGTCTGTGACCAAGACTACAAATCACCAAGTCTCTTCCACATGCTCTACATGCT 207
DB      267 CAGTCTGTGACCAAGACTACAAATCACCAAGTCTCTTCCACATGCTCTACATGCT 326
QY      208 CCTGTTTTTTGTTGGACTATATCAAAATGCGCTGGCGATGAGAGATTTCTTCAATCG 267
DB      327 CCTGTTTTTTGTTGGACTATATCAAAATGCGCTGGCGATGAGAGATTTCTTCAATCG 386
QY      268 GAGTAAATCAAACTTATTTCTTAAAGACACAGTATTTCTATCTCTCAATGAT 327
DB      387 GAGTAAATCAAACTTATTTCTTAAAGACACAGTATTTCTATCTCTCAATGAT 446
QY      328 TCTGACTTTTCATTCAAATTTCTAGTAGCAAACTGGAAACAGACACTGAGAAC 387

```

```

DB      447 TCTGACTTTTCATTCAAATTTCTTACTGATGCCAAACCTGGGAACAGACCACTGAGAAC 506
QY      388 TTTTGTGTGCAAGTACCTACCTCCGTCATATTTTATTTGACAAATGATATCAGTATTTCAAT 447
DB      507 TTTTGTGTGCAAGTACCTACCTCCGTCATATTTTATTTGACAAATGATATCAGTATTTCAAT 566
QY      448 CTTGGGACTGATAACTATCGATGCTTACCAAGAACCAACAGGCAATTTAAACATCCAA 507
DB      567 CTTGGGACTGATAACTATCGATGCTTACCAAGAACCAACAGGCAATTTAAACATCCAA 626
QY      508 CCCCCAAAATCTCTTGGGGGCTAAGATTTCTCTGTTGTCATCTGGGCAATTCATGTTCTT 567
DB      627 CCCCCAAAATCTCTTGGGGGCTAAGATTTCTCTGTTGTCATCTGGGCAATTCATGTTCTT 686
QY      568 ACTCTCTTTGCGTAACATGATTTCTGACCAACAGGCAAGCCAGACAAAGATGGAAGAA 627
DB      687 ACTCTCTTTGCGTAACATGATTTCTGACCAACAGGCAAGCCAGACAAAGATGGAAGAA 746
QY      628 ATGCTCTTTCTTAAATCAAGATTCGCTTATGCTGCAATGAAATAGTAAATTCATCTG 687
DB      747 ATGCTCTTTCTTAAATCAAGATTCGCTTATGCTGCAATGAAATAGTAAATTCATCTG 806
QY      688 TCAAGTCAATTTCTGATTAATTTCTTAATTTGATTTATGATATACATCAATACAAA 747
DB      807 TCAAGTCAATTTCTGATTAATTTCTTAATTTGATTTATGATATACATCAATACAAA 866
QY      748 AGAAGTGTACCGGTATATATGTAAGAACAGAGGGGTATAGTAAAGTCCCGAGAAAAGGT 807
DB      867 AGAAGTGTACCGGTATATATGTAAGAACAGAGGGGTATAGTAAAGTCCCGAGAAAAGGT 926
QY      808 GAAGGTAAAGTTTTCATTTATCATGCTGATATCTTATTTGTTTGTCTTCCATTT 867
DB      927 GAAGGTAAAGTTTTCATTTATCATGCTGATATCTTATTTGTTTGTCTTCCATTT 986
QY      868 TGCCCGAATTCCTTACACCTCGAGCCAAACCCGGATGCTTTGACAGCACTGCTGAAA 927
DB      987 TGCCCGAATTCCTTACACCTCGAGCCAAACCCGGATGCTTTGACAGCACTGCTGAAA 1046
QY      928 TACTCTGTTTATGTAAGAGAGACACTGTGTGTTAATCTCTTAAATGATCCCTGGA 987
DB      1047 TACTCTGTTTATGTAAGAGAGACACTGTGTGTTAATCTCTTAAATGATCCCTGGA 1106
QY      988 TCGTTCATCATTTTCTTCTTGGCAAGTCTTCAGAAATTTCTTGATATGATGCGAA 1047
DB      1107 TCGTTCATCATTTTCTTCTTGGCAAGTCTTCAGAAATTTCTTGATATGATGCGAA 1166
QY      1048 GTGCCCAATTTGCAACATCTCTGTCAGAGCAATAGAGAAAAAAGACAGATGCTGG 1107
DB      1167 GTGCCCAATTTGCAACATCTCTGTCAGAGCAATAGAGAAAAAAGACAGATGCTGG 1226
QY      1108 TGACCCCAATGGAAGAGACTCCCAATGTAACCAATTTAACTAAGCAATATTTCAATCTCT 1167
DB      1227 TGACCCCAATGGAAGAGACTCCCAATGTAACCAATTTAACTAAGCAATATTTCAATCTCT 1286
QY      1168 TGTGTTGAGAACCTGTTAAA 1187
DB      1287 TGTGTTGAGAACCTGTTAAA 1306

```

RESULT 4
 AAA27126
 ID AAA27126 standard; cDNA; 2286 BP.
 AC AAA27126;
 XX
 DT 21-AUG-2000 (first entry)
 XX
 DE Human 15625 receptor protein cDNA.
 XX
 KW Human; G-protein-coupled receptor; GPCR; 15625 receptor protein;
 KW glial cells; spleen; colon; liver; brain; T-cell; heart;
 KW red cell; thymus; B-cell; pancreas; disorder; chromosome 3;
 KW anaemia; neutropenia; thrombocytopenia; gene therapy; ss.

XX	Homo sapiens.
OS	
XX	
FH	Key
FT	Location/Qualifiers
CDS	261..1289
FM	/tag= a
FM	/*transl_except= (pos:297..299, aa:Asn)
FM	/product= "15625 receptor protein"
XX	
PN	M0200028028-A1.
PD	
PD	18-MAY-2000.
XX	
PF	05-NOV-1999; 99MO-US25956.
XX	
PR	06-NOV-1998; 98US-0187134.
PR	25-AUG-1999; 99US-0382918.
XX	
PA	(MILL-) MILLENNIUM PHARM INC.
PI	Gluckemann MA, Gu W, Welch NS;
DR	WPI: 2000-376543/32. P-PADB: AAI94444.
XX	
PS	Disclosure; Page 89-90; 97pp; English.
CC	The 15625 receptor protein is a novel G-coupled protein receptor (GPCR).
CC	The present sequence is the cDNA for this protein and was isolated by screening a human cDNA library with sequences homologous to other GPCRs.
CC	The 15625 receptor protein is expressed in the glial cells of the brain.
CC	It is also expressed in several other tissues. The 15625 receptor protein may be useful for producing antibodies which can be used to detect the presence of the receptor protein. The 15625 receptor polynucleotides are useful for generating probes, primers and antisense constructs. The polynucleotides encoding the 15625 receptor proteins can also be inserted into vectors to be used in gene therapy. The disorder that may be treated using the 15625 receptor protein polynucleotides and polypeptides include anaemia, neutropenia and thrombocytopenia.
CC	
XX	
SQ	Sequence 2286 BP; 786 A; 432 C; 353 G; 715 T; 0 other;
	Query Match 97.5%; Score 1156.8; DB 21; Length 2286; Best Local Similarity 99.8%; Pred. No. 1.4e-296; Matches 1158; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY	28 CTGCAAGCTGCATTAACACTTACTTACGTGATATTAATTAAACCCTCGAATCAACAGTT 87
Db	182 CTGCAAGCTGCATTAACACTTACTTACGTGATATTAATTAAACCCTCGAATCAACAGTT 241
OY	88 ATCAGGTAAACCAACAAGAATAATGCAAGCGGTGACAATTCACCTCTGGCGCTGGGAAC 147
Db	242 ATCAGGTAAACCAACAAGAATAATGCAAGCGGTGACAACCTCACCTCTGGCGCTGGGAAC 301
OY	148 CAGTCTGTGCACACAGAGACTACAAAATACCAAGTGCTCTCTCCCACTGCTTCACTGT 207
Db	302 CAGTCTGTGCACACAGAGACTACAAAATACCAAGTGCTCTCTCCCACTGCTTCACTGT 361
OY	208 CCTGTTTTTTTGTGACTTATCACAAAATGGCTGGCGATGGAGATTTTCTTCAAATCCG 267
Db	362 CCTGTTTTTTTGTGACTTATCACAAAATGGCTGGCGATGGAGATTTTCTTCAAATCCG 421
OY	268 GAGTAATCAAACTTATATTTTCTTAAAGAACAGTCATTTCTGATCTTTCATGAT 327
Db	422 GAGTAATCAAACTTATATTTTCTTAAAGAACAGTCATTTCTGATCTTTCATGAT 481
OY	328 TCTGACTTTTTCATTCAAAAATCTTAGATGCCAAACTGGGAACAGGACACATGAAAC 387
Db	482 TCTGACTTTTTCATTCAAAAATCTTAGATGCCAAACTGGGAACAGGACACATGAAAC 541

OY	388	TTTGTGTCGTAAGTACCTCCGCAATATTTATTTTCCAACTGATPACAGTATTCATTT	447
Db	542	TTTTGTGTGTAAGTACCTCCGCAATATTTATTTTCCAACTGATPACAGTATTCATTT	601
OY	448	CCTGGACTGATTAATCATGATTCCTACCGAAGACCAACGAGCCATTTAAACATCCAA	507
Db	602	CCTGGACTGATTAATCATGATTCCTACCGAAGACCAACGAGCCATTTAAACATCCAA	661
OY	508	CCCCAAAATCTCTTTGGGGGGCTAAGATTCTCTCTGTGTGATCTGGGCAATTCATGTTCTT	567
Db	662	CCCCAAAATCTCTTTGGGGGGCTAAGATTCTCTCTGTGTGATCTGGGCAATTCATGTTCTT	721
OY	568	ACTGCTTTGGCTTAACATGATTTGACCAACAGCAGCCGAGACACAAAGATGTGAAGA	627
Db	722	ACTGCTTTGGCTTAACATGATTTGACCAACAGCAGCCGAGACACAAAGATGTGAAGA	781
OY	628	ATGCTCTTTCCTTAATCAGAGTTCCGCTGTAGTCTGGCATGAATTAATTAATTCATCTG	687
Db	782	ATGCTCTTTCCTTAATCAGAGTTCCGCTGTAGTCTGGCATGAATTAATTAATTCATCTG	841
OY	688	TCAAGTCATTTTCGGATTAATTTCTTAATTTGTTATGTTATGTTATACACTCATTTACAA	747
Db	842	TCAAGTCATTTTCGGATTAATTTCTTAATTTGTTATGTTATGTTATACACTCATTTACAA	901
OY	748	AGAACTGACCGGTCATTCGTAAAGACGAGGGGTGATGTTAAATGCCCGAGAAAAAGTT	807
Db	902	AGAACTGACCGGTCATTCGTAAAGACGAGGGGTGATGTTAAATGCCCGAGAAAAAGTT	961
OY	808	GAAGCTCAAAAGTTTTCATTAATCATTTGCTGTATTTCTTTATTTGTTTGTCTCTTCCATTT	867
Db	962	GAAGCTCAAAAGTTTTCATTAATCATTTGCTGTATTTCTTTATTTGTTTGTCTCTTCCATTT	1021
OY	868	TGCCCGAATTTCTCTTAACCCCTGAGCCAAACCCGGGATGTTCTTTACGTGCACTGCTGA	927
Db	1022	TGCCCGAATTTCTCTTAACCCCTGAGCCAAACCCGGGATGTTCTTTACGTGCACTGCTGA	1081
OY	928	TACTCTGTTTATGTAAGAGAGACACCTGTGGTTAACTCTTAATGATCCTCGGA	987
Db	1082	TACTCTGTTTATGTAAGAGAGACACCTGTGGTTAACTCTTAATGATCCTCGGA	1141
OY	988	TCCGTTCAATGATTTTTCCTTTGCAAGTCCCTTCAGAAATTCCTTGATTAAGTATGCTGA	1047
Db	1142	TCCGTTCAATGATTTTTCCTTTGCAAGTCCCTTCAGAAATTCCTTGATTAAGTATGCTGA	1201
OY	1048	GTGCCCCAATTTCTGGAATCTGTCTGCCAGACAAATRGAAGAAAAAGACGATGGTGG	1107
Db	1202	GTGCCCCAATTTCTGGAATCTGTCTGCCAGACAAATRGAAGAAAAAGACGATGGTGG	1261
OY	1108	TGACCCCAATTAAGAGATCCCAATGTAACAAATTAACTAAGAAATATTTCAATCTCTT	1167
Db	1262	TGACCCCAATTAAGAGATCCCAATGTAACAAATTAACTAAGAAATATTTCAATCTCTT	1321
OY	1168	TGTGTTCAAGAACTCGTTAAA 1187	
Db	1322	TGTGTTCAAGAACTCGTTAAA 1341	
RESULT 5			
AAS98079			
ID AAS98079 standard; DNA; 1332 BP.			
AAS98079;			
12-MAR-2002 (first entry)			
Human DNA for potential G protein-coupled receptor #36.			
Human: G protein-coupled receptor; GPCR; ds; GATA; galanin receptor; Alzheimer's disease; amyotrophic lateral sclerosis; asthma; atherosclerosis; basal cell carcinoma; breast carcinoma; cardiomyopathy; chondrosarcoma; chronic obstructive pulmonary disease; Crohn's disease; depression; epilepsy; macular degeneration; lymphoma; melanoma;			

ID	Sequence	Location/Qualifiers
AA008695	standard; cDNA; 1330 BP.	
AA008695		
04-SEP-2001	(first entry)	
Human P2-purine receptor subtype, P2Y12 cDNA.		
Human: P2-purine receptor; P2Y12; cardiac; vasotropic; thrombolytic; cerebroprotective; gynaecological; ADP; adenosine 5'-diphosphate; angina myocardial infarction; ischemic attack; preclampsia; bleeding disorder; carotid endarterectomy; vascular graft surgery; brain disorder; migraine; vascular injury; schizophrenia; eating disorder; depression; angioplasty; peripheral vascular disease; platelet aggregation; restenosis; embolism; thrombocytopenic purpura; stroke; pertussis toxin-sensitive G protein; GI; disseminated intravascular coagulation; thrombosis; ss.		
Homo sapiens.		
Key	Location/Qualifiers	
CDS	73..1101	
FT	/*tag= a	
FT	/product= "Human P2-purine receptor, P2Y12"	
FT		
XX	MO200146454-A1.	
XX		
XX	28-JUN-2001.	
XX		
XX	26-DEC-2000; 2000MO-US34998.	
XX		
XX	23-DEC-1999; 99US-0171622.	
XX		
XX	(COR-) COR THERAPEUTICS INC.	
XX		
XX	Conley PB, Jantzen H, Ramakrishnan-Dubridge V, Julius DJ;	
XX	Hollinger G;	
XX	WPI; 2001-418082/44.	
XX	P-PSDB; AAE04386.	
XX		
XX	Novel isolated ADP receptor, termed P2Y12 receptor polypeptide, useful for identifying binding partners and for diagnostic applications -	
XX	Example 1; Page 85-87; 108bp; English.	
XX		
XX	The invention relates to ADP (adenosine 5'-diphosphate) receptor, termed as P2Y12 receptor and its corresponding cDNA molecule. P2Y12 receptor is the subtype of P2-purine receptor. The P2Y12 receptor is expressed selectively in the platelets and brain, and couples to a pertussis toxin sensitive G protein (Gi). P2Y12 receptor is a G protein-coupled receptor that responds to ADP. The invention also relates to a method for identifying an agent which is useful for modulating acute myocardial infarction, unstable angina, chronic stable angina, transient ischaemic attacks, strokes, peripheral vascular disease, preclampsia, deep venous thrombosis, embolism, disseminated intravascular coagulation, thrombotic thrombocytopenic purpura or a bleeding disorder; thrombotic and restenotic complications following angioplasty, carotid endarterectomy, post CABG (coronary artery bypass graft) surgery, vascular graft surgery, stent placements or insertion of endovascular devices and prostheses.	
XX	P2Y12 receptor is useful for identifying binding partners and for diagnostic applications. P2Y12 receptor provides targets for screening synthetic small molecules and combinatorial or naturally occurring compound libraries to regulate platelet aggregation, vascular injury, or disease as well as schizophrenia, eating disorders, depression, migraine and other brain disorders. The present cDNA sequence encodes human P2-purine receptor subtype, referred as P2Y12 receptor.	
XX		
XX	Sequence 1330 BP; 416 A; 289 C; 224 G; 401 T; 0 other;	

Query Match		97.0%: Score 1151.4: DB 22: Length 1330:	
Best Local Similarity		99.9%: Pred. No. 3e-295:	
Matches 1152: Conservative		0: Mismatches 1: Indels	0:
QY	CTGCATTAACACTACTACTACTGATATCATTTCAAACTCCAGAAATCAACAGTATACAGT	94	
Db	1 CTGCATTAACACTACTACTACTGATATCATTTCAAACTCCAGAAATCAACAGTATACAGT	60	
QY	95 AACCAACAAAGAAATGCAGCCGTCGACAAATCTCACCCTCTGGCCCTGGGAACACCACTCG	154	
Db	61 AACCAACAAAGAAATGCAGCCGTCGACAACTCACCCTCTGGCCCTGGGAACACCACTCG	120	
QY	155 TGCACACAGAGACTACAAATACACCAGGTCCTCTTCCCACTGCTGTACACTGCTCTGTTT	214	
Db	121 TGCACACAGAGACTACAAATACACCAGGTCCTCTTCCCACTGCTGTACACTGCTCTGTTT	180	
QY	215 TTTGTGGACCTTATCACAATATGGCCTGGCGATGAGAGATTTTCTTCAATCCGAGTAA	274	
Db	181 TTTGTGGACCTTATCACAATATGGCCTGGCGAGAGAGATTTTCTTCAATCCGAGTAA	240	
QY	275 TCAAACTTATATTTTCTTAAAGAACAGATTTTCTGATCTTCTCATTTCTGACT	334	
Db	241 TCAAACTTATATTTTCTTAAAGAACAGATTTTCTGATCTTCTCATTTCTGACT	300	
QY	335 TTTCCATTCAAAATTTCTAGTATGCCAAACTGGGAACAGACCACTGAGAACTTTGTG	394	
Db	301 TTTCCATTCAAAATTTCTAGTATGCCAAACTGGGAACAGACCACTGAGAACTTTGTG	360	
QY	395 TGTCAAGTACCTCCGCTATATTTATTTTCAAAATGATATATAGATTTTCACTCTGGGA	454	
Db	361 TGTCAAGTACCTCCGCTATATTTATTTTCAAAATGATATATAGATTTTCACTCTGGGA	420	
QY	455 CTGATTAATATCGATCGGTACACAGAAAGACACACAGGCACTTAAACATCCACCCCAA	514	
Db	421 CTGATTAATATCGATCGGTACACAGAAAGACACACAGGCACTTAAACATCCACCCCAA	480	
QY	515 AATCTCTTGGGGGCTAAGATTTCTCTGTGTGTCATCTGGGCAATTCACTGTTTACTCTCT	574	
Db	481 AATCTCTTGGGGGCTAAGATTTCTCTGTGTGTCATCTGGGCAATTCACTGTTTACTCTCT	540	
QY	575 TTGCCTAACATATTTCTGTGACCAACAGGCGCGAGAGACAAAGTGAAGAAATGCTCT	634	
Db	541 TTGCCTAACATATTTCTGTGACCAACAGGCGCGAGAGACAAAGTGAAGAAATGCTCT	600	
QY	635 TTCCTTAAATCAGAGTTCGGTCTAGTCTGGCATGAATATGAATTTACATCTGTCAAGTC	694	
Db	601 TTCCTTAAATCAGAGTTCGGTCTAGTCTGGCATGAATATGAATTTACATCTGTCAAGTC	660	
QY	695 AATTCTCGGATTAATTTCTTAATGTGTTATGTATATGATATACACTCATTTCAAAAGAACCTG	754	
Db	661 AATTCTCGGATTAATTTCTTAATGTGTTATGTATATGATATACACTCATTTCAAAAGAACCTG	720	
QY	755 TACCGGATATAGTAAAGACGAGGGGTGAGTAAAGTCCCAAGGAAAAGGTGAACCTG	814	
Db	721 TACCGGATATAGTAAAGACGAGGGGTGAGTAAAGTCCCAAGGAAAAGGTGAACCTG	780	
QY	815 AAAGTTTCATTTATCATTCGTATATCTTTATTTGTTTGTCTTCCATTTTGCCCGA	874	
Db	781 AAAGTTTCATTTATCATTCGTATATCTTTATTTGTTTGTCTTCCATTTTGCCCGA	840	
QY	875 ATTTCTTACACCTGAGCCAAACCCGGGGATGCTTTGATCTGGACATGCGTAAATATCTCG	934	
Db	841 ATTTCTTACACCTGAGCCAAACCCGGGGATGCTTTGATCTGGACATGCGTAAATATCTCG	900	
QY	935 TTTCTATGTGAAGAGAGCACTCTGTGTGTTAACTTCTTAAATGCATGCTGATCGCTTC	994	
Db	901 TTTCTATGTGAAGAGAGCACTCTGTGTGTTAACTTCTTAAATGCATGCTGATCGCTTC	960	
QY	995 ATCTATTTTTCCTTGGCAAGTCTTCAAGAAATTCCTTGATTAAGTATGCTGAAGGCCCC	1054	
Db	961 ATCTATTTTTCCTTGGCAAGTCTTCAAGAAATTCCTTGATTAAGTATGCTGAAGGCCCC	1020	
QY	1055 AATTCTGACACATCTCTGTCCAGAGACATAGAAAAAGAAACAGAGATGCTGTACCCA	1114	

|||||
Db 1021 AATCTGCAACATCTCTGTCTCCAGACAAATGAGAAAAAGACAGATGGTGGACCA 1080
QY 1115 AATGAGAGAGCTCCAAATGTAACAAATTAACTAAGAGAAATTTCAATCTCTTGTGTTTC 1174
Db 1081 AATGAGAGAGCTCCAAATGTAACAAATTAACTAAGAGAAATTTCAATCTCTTGTGTTTC 1140
QY 1175 AGAAGCTGCTTAAA 1187
Db 1141 AGAAGCTGCTTAAA 1153
RESULT 7
AAD08700
ID AAD08700 standard; cDNA; 1328 BP.
AC AAD08700;
XX
DT 04-SEP-2001 (first entry)
XX
DE Human P2-purineric receptor, P2Y12 truncated allelic variant cDNA.
XX
XX Human; P2-purineric receptor; P2Y12; cardiant; vasotrophic; thrombolytic;
KM cerebroprotective; gynaecological; ADP; adenosine 5'-diphosphate; angina;
KM myocardial infarction; ischaemic attack; pre-eclampsia; bleeding disorder;
KM carotid endarterectomy; vascular graft surgery; brain disorder; migraine;
KM vascular injury; schizophrenia; eating disorder; depression; angioplasty;
KM peripheral vascular disease; platelet aggregation; restenotic; embolism;
KM thrombocytopenic purpura; stroke; pertussis toxin-sensitive G protein;
KM G1; disseminated intravascular coagulation; thrombosis; mutant; mutein;
KM variant; ss.
XX
XX Homo sapiens.
OS Synthetic.
OS
XX
FH Key Location/Qualifiers
FT CDS 73..876
FT /tag= a
FT /product= "Human P2-purineric receptor P2Y12 mutant"
XX
XX WC200146454-A1.
XX
XX 28-JUN-2001.
XX
XX 26-DEC-2000; 2000MO-US34998.
XX
XX 23-DEC-1999; 9905-0171622.
XX
XX (COR-) COR THERAPEUTICS INC.
XX
XX Conley PB, Jantzen H, Ramakrishnan-Dubridge V, Julius DJ;
PI Holloper G;
XX
XX WPI; 2001-418082/44.
DR P-PSDB; AAE04387.
XX
XX Novel isolated ADP receptor, termed P2Y12 receptor polypeptide, useful
PT for identifying binding partners and for diagnostic applications
XX
XX Claim 19; page 90-91; 108bp; English.
XX
XX The invention relates to ADP (adenosine 5'-diphosphate) receptor, termed
CC as P2Y12 receptor and its corresponding cDNA molecule. P2Y12 receptor is
CC the subtype of P2-purineric receptor. The P2Y12 receptor is expressed
CC selectively in the platelets and brain, and couples to a pertussis toxin-
CC sensitive G protein (G1). P2Y12 receptor is a G protein-coupled receptor
CC that responds to ADP. The invention also relates to a method for
CC identifying an agent which is useful for modulating acute myocardial
CC infarction, unstable angina, chronic stable angina, transient ischaemic
CC attacks, strokes, peripheral vascular disease, pre-eclampsia, deep venous
CC thrombosis, embolism, disseminated intravascular coagulation, thrombotic
CC thrombocytopenic purpura or a bleeding disorder; thrombotic and
CC restenotic complications following angioplasty, carotid endarterectomy,

CC post CABG (coronary artery bypass graft) surgery, vascular graft surgery,
CC stent placements or insertion of endovascular devices and prostheses.
CC P2Y12 receptor is useful for identifying binding partners and for
CC diagnostic applications. P2Y12 receptor provides targets for screening
CC synthetic small molecules and combinatorial or naturally occurring
CC compound libraries to regulate platelet aggregation, vascular injury, or
CC disease as well as schizophrenia, eating disorders, depression, migraine
CC and other brain disorders. The present cDNA sequence encodes human P2-
CC purineric receptor subtype, P2Y12 truncated allelic variant. This
CC variant cDNA is obtained by deleting two bases 'CA' from position 769-790
CC of the wildtype P2-purineric receptor, P2Y12 cDNA.
XX
SQ Sequence 1328 BP; 415 A; 288 C; 224 G; 401 T; 0 other;
Query Match 95.8%; Score 1137.4; DB 22; Length 1328;
Best local Similarity 99.7%; Pred. No. 1,5e-291;
Matches 1150; Conservative 0; Mismatches 1; Indels 2; Gaps 1;
QY 35 CTGCAATTAATCTACTTACTGTGATGATTCATTAACCTCCAGAAATCAAGTTATCAGGT 94
Db 1 CTGCAATTAATCTACTTACTGTGATGATTCATTAACCTCCAGAAATCAAGTTATCAGGT 60
QY 95 AACCAAGAGAAATGCAAGCGGTGACAAATCTGACCTGCGCTGGGAGACACAGCTG 154
Db 61 AACCAAGAGAAATGCAAGCGGTGACAAATCTGACCTGCGCTGGGAGACACAGCTG 120
QY 155 TGCACAGAGACTACAAATACCAACAGGCTCTTCCACCTGCTACACTGCTGTTT 214
Db 121 TGCACAGAGACTACAAATACCAACAGGCTCTTCCACCTGCTACACTGCTGTTT 180
QY 215 TTGTGTGACTTATACAAATGGCTGCGATGAGATTTCTTCAATCCGAGTAAA 274
Db 181 TTGTGTGACTTATACAAATGGCTGCGATGAGATTTCTTCAATCCGAGTAAA 240
QY 275 TCAAACTTTATATTTTCTTAAAGAACACAGTCAATTCGATCTTCATGATTTTGACT 334
Db 241 TCAAACTTTATATTTTCTTAAAGAACACAGTCAATTCGATCTTCATGATTTTGACT 300
QY 335 TTTTCATTTCAAAATTTCTTGTGATGCCAACTGGGAAACAGACACTGAGAACTTTTGTG 394
Db 301 TTTTCATTTCAAAATTTCTTGTGATGCCAACTGGGAAACAGACACTGAGAACTTTTGTG 360
QY 395 TGTCAAGTTACCTCCGATATTTTATTTTCAATGATATTCAGTATTTTCTGAGGA 434
Db 361 TGTCAAGTTACCTCCGATATTTTATTTTCAATGATATTCAGTATTTTCTGAGGA 420
QY 455 CTGATACTATGATGCTGCTACCAAGAGACACAGGCGATTTAAACATCAACCCCAA 514
Db 421 CTGATACTATGATGCTGCTACCAAGAGACACAGGCGATTTAAACATCAACCCCAA 480
QY 515 AATCTCTTGGGGCTAAGATTTCTCTGTTGTGATCTGGGCATTCATGTTCTTACTCTCT 574
Db 481 AATCTCTTGGGGCTAAGATTTCTCTGTTGTGATCTGGGCATTCATGTTCTTACTCTCT 540
QY 575 TTGCGTAAACATGATTTGACCAACAGGACAGGACAGAGACAAAGATGGAAGAAATGCTCT 634
Db 541 TTGCGTAAACATGATTTGACCAACAGGACAGGACAGAGACAAAGATGGAAGAAATGCTCT 600
QY 635 TTGCGTAAACATGATTTGCTAGTCTGATGCAATTAATTAATTAATTAATTAATTAATTA 694
Db 601 TTGCGTAAACATGATTTGCTAGTCTGATGCAATTAATTAATTAATTAATTAATTAATTA 660
QY 695 ATTTTCTGATTAATTTCTTAATTTGTTATGTTATACACTCAATTAACAAAGAACTG 754
Db 661 ATTTTCTGATTAATTTCTTAATTTGTTATGTTATTAATTAATTAATTAATTAATTAATTA 720
QY 755 TACCGGTCATGCTAAGAGAGAGGCTGATGTTAAAGTCCCAAGAAAGGTAACGTC 814
Db 721 TACCGGTCATGCTAAGAGAGAGGCTGATGTTAAAGTCCCAAGAAAGGTAACGTC 780
QY 815 AAGGTTTCTATTCATGCTGATTCCTTATTTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 874
Db 781 AAGGTTTCTATTCATGCTGATTCCTTATTTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 838


```
QY 991 GTTCATCTATTTTCTCTTGGCAAGTCCTTCAGAAATTCCTTGATGAATGCTGAAGTG 1050
    |||||||
Db 903 GTTCATCTATTTTCTCTTGGCAAGTCCTTCAGAAATTCCTTGATGAATGCTGAAGTG 962
QY 1051 CCCCAATTCGCAACATCTCTGTCCCGAGACAAATGAGAAAAAACAAGATGCTGTGA 1110
    |||||||
Db 963 CCCCAATTCGCAACATCTCTGTCCCGAGACAAATGAGAAAAAACAAGATGCTGTGA 1022
QY 1111 CCCCAATGAGAGACTCCAAATGTAACAAATTAAGAAATATTTCAATCTCTTGT 1170
    |||||||
Db 1023 CCCCAATGAGAGACTCCAAATGTAACAAATTAAGAAATATTTCAATCTCTTGT 1082
QY 1171 GTTCAGAACTCGTTAAA 1187
    |||||||
Db 1083 GTTCAGAACTCGTTAAA 1099

RESULT 10
ABK50286
ID ABK50286 standard; DNA: 1143 BP.
AC ABK50286;
XX
XX 15-JUL-2002 (first entry)
DE Human DNA encoding the ADP-glucose receptor.
XX
XX Human; ds: gene; ADP-glucose receptor; G protein-coupled receptor;
XX agonist; antagonist; cardiovascular function disorder; vasorelaxation;
XX ischemia; angina pectoris; gastrointestinal disorder; diarrhoea;
XX immune disorder; immunodeficiency disorder; autoimmune disorder;
XX rheumatoid arthritis; bacterial infection; viral infection;
XX fungal infection; protozoal infection; respiratory disorder; asthma;
XX kidney disorder; glomerulonephritis; hepatobiliary disorder; cirrhosis;
XX endocrine disorder; adrenal dysfunction; musculoskeletal disorder;
XX osteoporosis; nervous system disorder; Alzheimer's disease;
XX psychotic disorder; depression; cancer; pain; glycogen storage disease;
XX disorder of body weight; AIDS; acquired immunodeficiency syndrome;
XX chromosome 3; Usher's syndrome type 3.
XX
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT CDS 52..1080
FT /*tag= a
FT /product= "ADP-glucose receptor"
XX
XX W0200224942-A2.
XX
XX PD 28-MAR-2002.
XX
XX PF 19-SEP-2001; 2001WO-US29523.
XX
XX PR 20-SEP-2000; 2000US-234025P.
XX PR 09-FEB-2001; 2001US-0780576.
XX
XX (REGC ) UNIV CALIFORNIA.
XX
XX PA Clivelli O, Nothacker H, Wang Z, Reinscheld R;
XX
XX PI WPI; 2002-383195/41.
XX
XX DR P-PSDB; AA080164.
XX
XX PT Identifying an agonist, antagonist or ligand of an ADP-glucose
XX receptor, for treating cardiovascular, gastrointestinal, kidney,
XX endocrine, immune disorders, and bacterial, viral, protozoal or fungal
XX infections
XX
XX PS Example 1: Fig 1; 86pp; English.
XX
XX CC The invention relates to identifying an ADP-glucose (ADP-G) receptor (a
XX G protein-coupled receptor) agonist, antagonist or ligand, comprises
```

```
CC contacting the receptor with one or more candidate compounds so that the
CC receptor produces a G protein-coupled signal in response to ADP-G or
CC selectively binds ADP-G; and identifying the candidate molecule that
CC alters signal production as an agonist, antagonist or binds as a ligand.
CC Also included are altering signalling through an ADP-G receptor, by
CC contacting a cell expressing the receptor with ADP-G or the ant/agonist,
CC ameliorating an ADP-G receptor associated condition, by administering a
CC therapeutic composition comprising ADP-G or the ant/agonist to an
CC individual and a composition comprising the ADP-G receptor and ADP-G.
CC ADP-G or the ant/agonist is useful for treating an ADP-G receptor
CC associated condition e.g. cardiovascular function disorder, where the
CC therapeutic composition induces vasorelaxation. The new methods are
CC useful in identifying ant/agonists and ligands of the receptor.
CC The ant/agonists are useful therapeutically for preventing or
CC ameliorating conditions associated with the receptor such as
CC cardiovascular disorders (e.g. ischemia, hypertension, hypotension,
CC angina pectoris, myocardial infarction, stroke, congestive heart
CC failure, shock, erectile dysfunction, orthostatic intolerance and
CC migraine), gastrointestinal disorders (e.g. diarrhoea, gastritis,
CC inflammatory bowel disease), immune disorders (e.g. immunodeficiency
CC disorders, autoimmune disorders, rheumatoid arthritis), infections
CC caused by bacteria, fungi, protozoa or virus, respiratory disorders
CC (e.g. asthma, pneumonia, bronchitis), kidney disorders (e.g.
CC glomerulonephritis, renal failure, lupus), hepatobiliary disorders
CC (e.g. jaundice, cirrhosis, hepatitis), endocrine disorders (e.g.
CC pituitary, thyroid or adrenal dysfunctions), musculoskeletal disorders
CC (e.g. osteoporosis, muscular dystrophies), nervous system disorders
CC (e.g. Parkinson's and Alzheimer's disease), psychotic disorders (e.g.
CC depression, anxiety, schizophrenia), pain, glycogen storage
CC diseases and disorders of body weight (e.g. as a result of cancer
CC or AIDS, acquired immunodeficiency syndrome). The gene for the ADP-G
CC receptor is located on chromosome 3 in a region associated with Usher
CC syndrome type 3 (progressive hearing loss). The ligand is useful
CC therapeutically, in detecting normal or abnormal expression of the
CC receptor in an isolated sample or in in vivo diagnostic imaging
CC procedures, and targeting specifically a diagnostic group to cells and
CC tissues that express the receptor. The present sequence encodes the
XX ADP-G receptor.
XX
XX SQ Sequence 1143 BP; 326 A; 257 C; 201 G; 359 T; 0 other;

Query Match 92.1%; Score 1093.8; DB 24; Length 1143;
Best Local Similarity 99.8%; Pred. No. 5.4e-280;
Matches 1095; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 91 AGGTACCAACAGAAATGCAAGCGGTGACAAATCTCCTCGGCTGGACACACAG 150
    |||||||
Db 36 AGGTACCAACAGAAATGCAAGCGGTGACAAATCTCCTCGGCTGGACACACAG 95
QY 151 TCTGTGCACGAGACTACAAATCACCAGTCCCTCTCCACTGCTCTACACTGTCT 210
    |||||||
Db 96 TCTGTGCACGAGACTACAAATCACCAGTCCCTCTCCACTGCTCTACACTGTCT 155
QY 211 GTTTTGTGTGACTATATCAAAATGGCGCTGGCATAGAGATTTTCTTCAATCCGAG 270
    |||||||
Db 156 GTTTTGTGTGACTATATCAAAATGGCGCTGGCATAGAGATTTTCTTCAATCCGAG 215
QY 271 TAAATCAAACTTTATTTATTTCTTAAGAACACAGTCAATTCGTATCTTCAGATGCT 330
    |||||||
Db 216 TAAATCAAACTTTATTTATTTCTTAAGAACACAGTCAATTCGTATCTTCAGATGCT 275
QY 331 GACTTTTCCATTCAAAATCTTAGTAGATGCCAAACTGGGAACAGACACAGAACTTT 390
    |||||||
Db 276 GACTTTTCCATTCAAAATCTTAGTAGATGCCAAACTGGGAACAGACACAGAACTTT 335
QY 391 TGTGTGCAAAATTCACCGCGCATATTTTATTTGACAAATGTAATACATATTTCTTCT 450
    |||||||
Db 336 TGTGTGCAAAATTCACCGCGCATATTTTATTTGACAAATGTAATACATATTTCTTCT 395
QY 451 GGAGCTGATTAATTCGATCGCTACAGAAAGACACACAGGCAATTTAAACATCCACCC 510
    |||||||
Db 396 GGAGCTGATTAATTCGATCGCTACAGAAAGACACACAGGCAATTTAAACATCCACCC 455
```

```
QY 511 CAAAAATCTCTGGGGGCTAAGATTCTCTCTGTGTGATCATCTGGGCAATCAATGTTCTTACT 570
    |||
Db 456 CAAAAATCTCTGGGGGCTAAGATTCTCTCTGTGTGATCATCTGGGCAATCAATGTTCTTACT 515
QY 571 CTTTGGCCATCATGATCTCTGACCAAGAGGAGCGGAGAGAGCAAAATGTAAGAATG 630
    |||
Db 516 CTTTGGCCATCATGATCTCTGACCAAGAGGAGCGGAGAGAGCAAAATGTAAGAATG 575
QY 631 CTTTGGCCATCAATCAAGAGTGGTGTAGTCTGTGGCATGAAATAGTAATATCATCTGTCA 690
    |||
Db 576 CTTTGGCCATCAATCAAGAGTGGTGTAGTCTGTGGCATGAAATAGTAATATCATCTGTCA 635
QY 691 AGCATATTTCTGGATTAATTTCTTAATGTATGTATGTATGATACACTATTCAAAAGA 750
    |||
Db 636 AGCATATTTCTGGATTAATTTCTTAATGTATGTATGATGATTAACACTATTCAAAAGA 695
QY 751 ACTGACCGGTCAATGATGAAGAGAGGGGTGATGATTAAGTCCCAAGGAAAAAGGTGA 810
    |||
Db 696 ACTGACCGGTCAATGATGAAGAGAGGGGTGATGATTAAGTCCCAAGGAAAAAGGTGA 755
QY 811 CGTCAAAAGTTTCATATGATGCTGTATCTTATTTGTTTGTCTCTTCCATTTTGC 870
    |||
Db 756 CGTCAAAAGTTTCATATGATGCTGTATCTTATTTGTTTGTCTCTTCCATTTTGC 815
QY 871 CGCAATTCCTTACACCTGAGCCAAACCGGGGATGCTTTGACTGCACGCTGAAAATAC 930
    |||
Db 816 CGCAATTCCTTACACCTGAGCCAAACCGGGGATGCTTTGACTGCACGCTGAAAATAC 875
QY 931 TCTGTCTATGTAAGAGAGACGCTGTGTGTTAACTCTTAAATGACATGCTGTGATCC 990
    |||
Db 876 TCTGTCTATGTAAGAGAGACGCTGTGTGTTAACTCTTAAATGACATGCTGTGATCC 935
QY 991 GTTCATCTATTTTTCCTTTCGCAAGTCCTTCAGAAATTCCTTGATTAAGTGTGAAAGTG 1050
    |||
Db 936 GTTCATCTATTTTTCCTTTCGCAAGTCCTTCAGAAATTCCTTGATTAAGTGTGAAAGTG 995
QY 1051 CCCCAATTCGCAACATCTCTGTCCAGAGCAATAGAGAAAAAGAACAGAGATGTTGTA 1110
    |||
Db 996 CCCCAATTCGCAACATCTCTGTCCAGAGCAATAGAGAAAAAGAACAGAGATGTTGTA 1055
QY 1111 CCCAAATGAAGAGACTCAATGTAACAATTAAGTAAGAAATATTTCAATCTCTTGT 1170
    |||
Db 1056 CCCAAATGAAGAGACTCAATGTAACAATTAAGTAAGAAATATTTCAATCTCTTGT 1115
QY 1171 GTTCAGACTCGTTAAA 1187
    |||
Db 1116 GTTCAGACTCGTTAAA 1132

RESULT 11
AA168802
ID AA168802 standard; DNA; 2245 BP.
XX
AC AA168802;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human PAFR3 DNA.
XX
KW PAFR-3; platelet activating factor receptor; human; chromosome 3;
KW thrombocyte activation; hypotension; vascular permeability;
KW bronchoconstriction; transgenic animal; knockout animal; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 467..1145
FT /tag= a
FT /product= "PAFR3"
XX
PN DE10020073-A1.
XX
PD 25-OCT-2001.
```

```
XX 22-APR-2000; 2000DE-1020073.
PF 22-APR-2000; 2000DE-1020073.
XX
PR 22-APR-2000; 2000DE-1020073.
XX
PA (BRUE/) BRUESS M.
PA (BOEN/) BOENISCH H.
XX
PI Brues M, Boenisch H;
XX
DR WPI: 2002-027296/04.
XX
PT P-PDB; AAG80236.
XX
PS New human platelet-activating factor (PAF) receptor-3 gene, useful for
disclosure: Page 3-4; 6pp; German.
XX
CC This invention describes a novel human platelet-activating factor (PAF)
CC receptor-3 (PAFR-3) gene (I). The protein encoded by (I) probably
CC modulates the activity of PAF, which is involved in many
CC (patho)physiological processes, e.g., thrombocyte activation,
CC hypotension, increased vascular permeability, bronchoconstriction etc.
CC (I), and derived (anti)sense oligonucleotides, are useful in the
CC treatment and diagnosis of (I)-related diseases; for producing
CC transgenic/knockout animals, and for recombinant expression of the
CC protein (II) that it encodes. (II) is useful in ligand-binding studies
CC and screening assays, also for treatment and diagnosis of (II)-related
CC diseases. This sequence encodes the human leucocyte platelet-activating
CC factor receptor (PAFR-3) protein described in the invention.
XX
SQ Sequence 2245 BP; 676 A; 466 C; 397 G; 706 T; 0 other;
Query Match 92.1%; Score 1093.8; DB 24; Length 2245;
Best Local Similarity 99.8%; Pred No. 7e-280;
Matches 1095; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 91 AGGTAACAACAAGAAATGCAAGCCGTCGACAAATCTCACCTGCGCTGGAGAACACAG 150
    |||
Db 451 AGGTAACAACAAGAAATGCAAGCCGTCGACAAATCTCACCTGCGCTGGAGAACACAG 150
QY 151 TCTGTGACACAGAGACTACAAAATACACAGGTCCTTCCACGCTCTACTACTGCT 210
    |||
Db 511 TCTGTGACACAGAGACTACAAAATACACAGGTCCTTCCACGCTCTACTACTGCT 210
QY 211 GTTTTGTGACATATCAAAATGGCTGGGATGAGAGATTTCTTCAATCCGAG 270
    |||
Db 571 GTTTTGTGACATATCAAAATGGCTGGGATGAGAGATTTCTTCAATCCGAG 270
QY 271 TAAATCAAACTTATTTTCTTAAAGACACAGTCATTTCTGATCTTCATGATTC 330
    |||
Db 631 TAAATCAAACTTATTTTCTTAAAGACACAGTCATTTCTGATCTTCATGATTC 330
QY 331 GACTTTTCATTCAAAATTTCTAGTATGCCAAATGCGGAGAACAGGACCTGAACTTT 390
    |||
Db 691 GACTTTTCATTCAAAATTTCTAGTATGCCAAATGCGGAGAACAGGACCTGAACTTT 390
QY 391 TGTGTCAAGTACCTCGCTCATATTTTATTTTACAAATGATATAGATTTCAATTC 450
    |||
Db 751 TGTGTCAAGTACCTCGCTCATATTTTATTTTACAAATGATATAGATTTCAATTC 450
QY 451 GGGAGTATACATGATGATGCTGACAGAAAGACACAGGCAATTTAAACATCCAAACC 510
    |||
Db 811 GGGAGTATACATGATGATGCTGACAGAAAGACACAGGCAATTTAAACATCCAAACC 510
QY 511 CAAAAATCTCTGGGGGCTAAGATTCTCTCTGTGTGATCATCTGGGCAATCAATGTTCTTACT 570
    |||
Db 871 CAAAAATCTCTGGGGGCTAAGATTCTCTCTGTGTGATCATCTGGGCAATCAATGTTCTTACT 570
QY 571 CTTTGGCCATCATGATCTCTGACCAAGAGGAGCGGAGAGAGCAAAATGTAAGAATG 630
    |||
Db 931 CTTTGGCCATCATGATCTCTGACCAAGAGGAGCGGAGAGAGCAAAATGTAAGAATG 630
```

PR	28-MAY-1999;	99US-0136436.
PR	28-MAY-1999;	99US-0136437.
PR	28-MAY-1999;	99US-0136439.
PR	28-MAY-1999;	99US-0136567.
PR	28-MAY-1999;	99US-0137127.
PR	28-MAY-1999;	99US-0137311.
PR	28-MAY-1999;	99US-0141448.
PR	29-SEP-1999;	99US-0146555.
PR	29-SEP-1999;	99US-0156633.
PR	29-SEP-1999;	99US-0156634.
PR	29-SEP-1999;	99US-0156653.
PR	01-OCT-1999;	99US-0157280.
PR	01-OCT-1999;	99US-0157281.
PR	01-OCT-1999;	99US-0157282.
PR	01-OCT-1999;	99US-0157293.
PR	01-OCT-1999;	99US-0157294.
PR	12-OCT-1999;	99US-0416760.
PR	12-OCT-1999;	99US-0417044.
PR		
PA	(AREN-) ARENA PHARM INC.	
PI		
PI	Chen R, Dang HT, Liaw CW, Lin I;	
XX		
XX	WPI: 2000-400068/34.	
DR	P-PSDB: MAY71306.	
XX		
PT	Novel human orphan G protein-coupled receptors and the encoding cDNAs	
PT	for use in the identification of G protein-coupled/receptor agonists -	
XX		
ES	Claim 61; Page 81-82; 102pp; English.	
XX		
CC	The present sequence is a cDNA encoding hCHN8, an endogenous human	
CC	orphan G protein-coupled receptor (GPCR), expressed in left and right	
CC	cerebellum, kidney and lung. The hCHN8 cDNA was identified using full	
CC	length EST (expressed sequence tag) 764455 as a probe.	
CC	The orphan GPCR of the invention, like all GPCRs has seven transmembrane	
CC	alpha helices with an extracellular N-terminus and an intracellular	
CC	C-terminus. However, no endogenous ligands has yet been identified for	
CC	the proteins of the invention. The orphan GPCRs may be used in the	
CC	identification of their endogenous ligands, and to screen potential GPCR	
CC	agonists and antagonists for use as pharmaceutical agents. The proteins	
CC	may also be used in the study of GPCR-mediated signalling cascades, and	
CC	to elucidate their precise role in normal and diseased human conditions.	
CC	Nucleic acid encoding human orphan GPCRs may be used for tissue	
CC	localisation expression analysis to provide information about their	
CC	function in healthy and pathological states.	
XX		
SO	Sequence 1029 BP; 285 A; 233 C; 186 G; 325 T; 0 other:	
	Query Match	86.7%; Score 1029; DB 21; Length 1029;
	Best Local Similarity	100.0%; Pred. No. 7.9e-263;
	Matches 1029; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	107 ATGCAAGCCGCTGCACAAATCTCACCTCTGCGCTGGAGACCAACGATCTGTGCACGAGAGC	166
DB	1 ATGCAAGCCGCTGCACAAATCTCACCTCTGCGCTGGAGACCAACGATCTGTGCACGAGAGC	60
OY	167 TACAACATCCACGAGTCTCTCCACAGCGCTCTACACGTCCTGTTTCTTGAGACTT	226
DB	61 TACAACATCCACGAGTCTCTCCACAGCGCTCTACACGTCCTGTTTCTTGAGACTT	120
OY	227 ATCACAATATGCGCTGGCGATGAGATTTCTTCAATCCGAGTAATCAACATTTATT	286
DB	121 ATCACAATATGCGCTGGCGATGAGATTTCTTCAATCCGAGTAATCAACATTTATT	180
OY	287 ATTTTCTTAAGACACACTATTTCTGATCTTCATGATGATTTGACTTTCATTCANA	346
DB	181 ATTTTCTTAAGACACACTATTTCTGATCTTCATGATGATTTGACTTTCATTCANA	240
OY	347 ATTTCTAGTGAAGCAAACTGGGAACAGACACACGAAAGAACTTTGTGTCAAGTTACC	406
DB	241 ATTTCTAGTGAAGCAAACTGGGAACAGACACACGAAAGAACTTTGTGTCAAGTTACC	300


```

Db 241 ATCTTACTGATGCGCAACCTGGGAACAGACACCTGAAACCTTTTGTCTCAACTTAC 300
Qy 407 TCCGTCATATTTTATTTTACAAATGATATATCATTTTCCTGCGGAGCTGATACTATC 466
Db 301 TCCGTCATATTTTATTTTACAAATGATATATCATTTTCCTGCGGAGCTGATACTATC 360
Qy 467 GATCGCTACAGAGAACACAGCCAGCCATTTAAACATCCAAACCCAAAATCTCTTGGGG 526
Db 361 GATCGCTACAGAGAACACAGCCAGCCATTTAAACATCCAAACCCAAAATCTCTTGGGG 420
Qy 527 GCTAAGATTTCTCTCTGTGTGATCTGGGCAATTCATGTTCTTACTCTCTTGGCTTAACATG 586
Db 421 GCTAAGATTTCTCTCTGTGTGATCTGGGCAATTCATGTTCTTACTCTCTTGGCTTAACATG 480
Qy 587 ATTCTGACCAAGAGCAGCCAGAGACAAAGATGGAAGAAATGCTTTCTTAAATCA 646
Db 481 ATTCTGACCAAGAGCAGCCAGAGACAAAGATGGAAGAAATGCTTTCTTAAATCA 540
Qy 647 GAGTTGGGTCTAGTCTGGCATGAATAGTAATACATCTGTCAAGTCATTTTCTGAT 706
Db 541 GAGTTGGGTCTAGTCTGGCATGAATAGTAATACATCTGTCAAGTCATTTTCTGAT 600
Qy 707 AATTCTTAATTTTATTTGATGTATATACACTCATTTACAAAAGAACTTACCGGTCAATC 766
Db 601 AATTCTTAATTTTATTTGATGTATATACACTCATTTACAAAAGAACTTACCGGTCAATC 660
Qy 767 GTAAGAACAGAGGGGTAGTAGTAAGTCCCGAGAAAGAGTAAGCTCAAGTTTCTT 826
Db 661 GTAAGAACAGAGGGGTAGTAGTAAGTCCCGAGAAAGAGTAAGCTCAAGTTTCTT 720
Qy 827 ATCATGCTGATTTCTTTATTTGTTTGTCTTTCATTTTGGCCGAATTCCTTACAC 886
Db 721 ATCATGCTGATTTCTTTATTTGTTTGTCTTTCATTTTGGCCGAATTCCTTACAC 780
Qy 887 CTGAGCCCAAAACCCGGAGTCTTTGACCTGACCTGTGAATACTCTGTTCTATGTGA 946
Db 781 CTGAGCCCAAAACCCGGAGTCTTTGACCTGACCTGTGAATACTCTGTTCTATGTGA 840
Qy 947 GAGAGCACTGTGTAACTTCTTAAATGATGATGCTGATGCGGTCATCATTTTTC 1006
Db 841 GAGAGCACTGTGTAACTTCTTAAATGATGATGCTGATGCGGTCATCATTTTTC 900
Qy 1007 CTTTGCAAGTCTTCAAGAAATCTGTAGTAAGTATGCTGAAGTGGCCCAATTCGCAACA 1066
Db 901 CTTTGCAAGTCTTCAAGAAATCTGTAGTAAGTATGCTGAAGTGGCCCAATTCGCAACA 960
Qy 1067 TCTCTGTCAGAGCAATAGGAAAAAGAAACAGATGCTGTGACCCCAATGAGAGACT 1126
Db 961 TCTCTGTCAGAGCAATAGGAAAAAGAAACAGATGCTGTGACCCCAATGAGAGACT 1020
Qy 1127 CCAATGTAA 1135
Db 1021 CCAATGTAA 1029

```

RESULT 14
ABLS9205
ID ABL59205 standard; DNA: 1029 BP.

AC ABL59205;
XX 27-SEP-2002 (first entry)
XX Nucleotide sequence of human ADP receptor P2YAC protein.
XX Human: ADP receptor; P2YAC; antiplatelet agent; thrombotic disease;
KW Ischaemic disease; ss.
XX Homo sapiens.
XX Key location/Qualifiers
FH CDS 1..1029

```

FT /*tag= a
FT /product= "ADP receptor P2YAC protein"
PN MO200236631-A1.
XX 10-MAY-2002.
PD 31-OCT-2001; 2001WO-JP09534.
XX 01-NOV-2000; 2000JP-0334721.
XX 11-JAN-2001; 2001JP-0003577.
PA (YAMA ) YAMANOUCHI PHARM CO LTD.
PI Takasaki J, Matsumoto M, Kamohara M, Saito T, Ohishi T, Soga T;
XX WPI: 2002-519236/55.
XX P-PSDB: ABB77868.
DR
PT Method for screening anti-platelet agents with human ADP receptor P2YAC
PT protein, its functional equivalent, or its homologous protein, as tool,
PT for treating e.g. thrombotic or ischaemic diseases
XX
XX Claim 1; Page 46-48; 56pp; Japanese.
XX
XX The present sequence encodes a human ADP receptor P2YAC protein. It is
XX used as a screening tool for screening for antiplatelet agents. The
XX method is used for screening for antiplatelet agents, which can then
XX be used for treating e.g. thrombotic or ischaemic diseases.
XX
XX Sequence 1029 BP; 285 A; 234 C; 186 G; 324 T; 0 other;
SQ
Query Match 86.6%; Score 1027.4; DB 24; Length 1029;
Best Local Similarity 99.9%; Pred. No. 2.1e-262;
Matches 1028; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 107 ATGCAAGCCGTGACAAATCTACCTCTGCGCCTGGGAACACGATCTGTGACAGAGAC 166
Db 1 ATGCAAGCCGTGACAAATCTACCTCTGCGCCTGGGAACACGATCTGTGACAGAGAC 60
Qy 167 TACAATAATCACCAGAGTCTCTCCCACTGCTCTACCTGCTCTGTTTCTTGTGACT 226
Db 61 TACAATAATCACCAGAGTCTCTCCCACTGCTCTACCTGCTCTGTTTCTTGTGACT 120
Qy 227 ATCAAAATGGCGTGGCATGAGATTTCTTCAATCCGAGTAATCAACTTAAAT 286
Db 121 ATCAAAATGGCGTGGCATGAGATTTCTTCAATCCGAGTAATCAACTTAAAT 180
Qy 287 ATTTTCTTAAGAACACAGTCTTCTGATCTTCTCATGATTTCTGACTTTTCCATTCAA 346
Db 181 ATTTTCTTAAGAACACAGTCTTCTGATCTTCTCATGATTTCTGACTTTTCCATTCAA 240
Qy 347 ATCTTAGTATGCCAACTGGGAACAGGACCACTGAGAACTTTGTGTGAAGTTAC 406
Db 241 ATCTTAGTATGCCAACTGGGAACAGGACCACTGAGAACTTTGTGTGAAGTTAC 300
Qy 407 TCCGTCATATTTTATTTTACAAATGATATATCATTTTCCTGCGGAGCTGATACTATC 466
Db 301 TCCGTCATATTTTATTTTACAAATGATATATCATTTTCCTGCGGAGCTGATACTATC 360
Qy 467 GATCGCTACAGAGAACACAGCCAGCCATTTAAACATCCAAACCCAAAATCTCTTGGGG 526
Db 361 GATCGCTACAGAGAACACAGCCAGCCATTTAAACATCCAAACCCAAAATCTCTTGGGG 420
Qy 527 GCTAAGATTTCTCTCTGTGTGATCTGGGCAATTCATGTTCTTACTCTCTTGGCTTAACATG 586
Db 421 GCTAAGATTTCTCTCTGTGTGATCTGGGCAATTCATGTTCTTACTCTCTTGGCTTAACATG 480
Qy 587 ATTCTGACCAAGAGCAGCCAGAGACAAAGATGGAAGAAATGCTTTCTTAAATCA 646
Db 481 ATTCTGACCAAGAGCAGCCAGAGACAAAGATGGAAGAAATGCTTTCTTAAATCA 540
Qy 647 GAGTTGGGTCTAGTCTGGCATGAATAGTAATACATCTGTCAAGTCATTTTCTGAT 706

```

Db 541 GAGTTGGGCTAGCTGGCATGAAATAGTAATATACATCTGTCAAGCATTTCTGTGATTT
 QY 707 AATTTCTTAATGTGTAATGTATGTTATATACATCATATACAAAGAACTGTACCGGTACATAC
 Db 601 AATTTCTTAATGTGTAATGTATGTTATATACATCATATACAAAGAACTGTACCGGTACATAC
 QY 767 GTAGAACAAGGGGTGTAGTAAAGTCCCAAGAAAAAGTGAACGTCAAAAGTTTTCATT
 Db 661 GTAGAACAAGGGGTGTAGTAAAGTCCCAAGAAAAAGTGAACGTCAAAAGTTTTCATT
 QY 827 ATATTCTGTATTTCTTATTTGTTTGTCTTCTTCCATTTTGGCCGAATTCCTTACACC
 Db 721 ATCATTTCTATTTCTTATTTGTTTGTCTTCTTCCATTTTGGCCGAATTCCTTACACC
 QY 887 CTGAGCCAAACCCGGGATGCTTTGACTGCATGCTGAAATATCTGTTCTATGTGAAA
 Db 781 CTGAGCCAAACCCGGGATGCTTTGACTGCATGCTGAAATATCTGTTCTATGTGAAA
 QY 947 GAGAGCACTGTGTGTTAACTTCTTAATGATGCTGATCCGTTCAATCTATTTTTC
 Db 841 GAGAGCACTGTGTGTTAACTTCTTAATGATGCTGATCCGTTCAATCTATTTTTC
 QY 1007 CTTTGCAAGTCTTCAAGAAATCTTGTATAGTATGCTGAAGTGGCCCAATTCGACACA
 Db 901 CTTTGCAAGTCTTCAAGAAATCTTGTATAGTATGCTGAAGTGGCCCAATTCGACACA
 QY 1067 TCTCTGTCCAGGACATAGGAAAAAGAAACAGAGATGGTGTGACCCCAATGAGAGACT
 Db 961 TCTCTGTCCAGGACATAGGAAAAAGAAACAGAGATGGTGTGACCCCAATGAGAGACT
 QY 1127 CCAATGTAA 1135
 Db 1021 CCAATGTAA 1029

RESULT 15

ABA98535 standard; cDNA; 1029 BP.

AC ABA98535;

DT 25-APR-2002 (first entry)

DE Human G protein-coupled receptor, IGPcR17, coding sequence.

XX Human; G protein-coupled receptor; IGPcR17; gene; analgesic; neuroleptic;

KM tranquilizer; antiparkinsonian; neuroprotective; nootropic;

KM anticonvulsant; metabolic; anorectic; anabolic; antiinflammatory;

KM antidiarrheic; osteopathic; antiasthmatic; antiallergic; antialtritic;

KM immunosuppressive; gene therapy; psychiatric disorder;

KM central nervous system disorder; movement dysfunction; schizophrenia;

KM multiple sclerosis; Alzheimer's disease; kidney disease; obesity;

KM gastrointestinal disorder; osteoporosis; infection;

KM gynecological disorder; ss.

OS Homo sapiens.

PD 10-JAN-2002.

PE 02-JUL-2001; 2001MO-EP07532.

PR 30-JUN-2000; 2000US-215759P.

PA (INGE-) INGENIUM PHARM AG.

PI Wattler F, Wattler S, Trommler P, Nehls WC;
 XX MPI; 2002-14080/18.
 DR P-PSDB; AAM48353.
 XX
 PT New human or mouse G protein-coupled receptor protein, IGPcR17, useful
 PT for diagnosis, prevention, amelioration or treatment of central nervous
 PT system disorders such as Tourette's syndrome, Parkinson's disease and
 PT pain
 PS Claim 1; Fig 1; 71pp; English.

CC The present sequence is the coding sequence for human G protein-coupled
 CC receptor (GPcR) protein, IGPcR17. The coding sequence for IGPcR17 is
 CC useful in gene therapy for prevention, amelioration or treatment of
 CC diseases characterised by aberrant expression or activity of IGPcR17,
 CC where the disease is a psychiatric or central nervous system (CNS)
 CC disorder associated with signal processing in CNS such as learning and
 CC memory disorders, movement dysfunctions, tics, tremor, Tourette's
 CC syndrome, Parkinson's disease, Huntington's disease, dyskinesias,
 CC dystonia, pain and spasms. In addition, IGPcR17 and its coding sequence
 CC are useful in diagnosis, prevention, amelioration or treatment of
 CC diseases associated with signal processing in CNS, schizophrenia,
 CC episodic paroxysmal anxiety (EPA) disorders such as obsessive compulsive
 CC disorder (COD), multiple sclerosis, Alzheimer's disease/dementia,
 CC anorexia, kidney diseases such as renal failure, obesity,
 CC gastrointestinal disorders such as irritable bowel syndrome (IBS),
 CC diarrhoea, motility disorders and conditions of delayed gastric emptying,
 CC osteoporosis, infections such as bacterial, fungal, protozoal and viral
 CC infections, asthma, allergy, arthritis, sepsis and gynecological
 CC disorders.

SQ Sequence 1029 BP; 285 A; 234 C; 186 G; 324 T; 0 other;

Query Match 86.6%; Score 1027.4; DB 24; Length 1029;

Best Local Similarity 99.9%; Pred. No. 2,1e-262;

Matches 1028; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 107 ATGCAGCCGCTGCACAAATCTACCTCTGCGCTGGGAACACAGTGTGCACAGAGAC 166
 Db 1 ATGCAGCCGCTGCACAAATCTACCTCTGCGCTGGGAACACAGTGTGCACAGAGAC 60
 QY 167 TACAAAATTCACCCAGGTCCTTCCACTGCTGTACACTGTCTGTTTGTGGACTT 226
 Db 61 TACAAAATTCACCCAGGTCCTTCCACTGCTGTACACTGTCTGTTTGTGGACTT 120
 QY 227 ATACAAATGCGCTGGGAGATGAGATTTCTTCAATTCGGAGTAAATCAATATT 286
 Db 121 ATACAAATGCGCTGGGAGATGAGATTTCTTCAATTCGGAGTAAATCAATATT 180
 QY 287 ATTTTCTTAAAGAACACAGTCATTTCTGATCTTCTCATGATTCGACTTTCATTCAAA 346
 Db 181 ATTTTCTTAAAGAACACAGTCATTTCTGATCTTCTCATGATTCGACTTTCATTCAAA 240
 QY 347 ATTTCTTATGATGCGCAATCTGGGAACAGACACTGAGAACTTTGTGTCAAGTTAC 406
 Db 241 ATTTCTTATGATGCGCAATCTGGGAACAGACACTGAGAACTTTGTGTCAAGTTAC 300
 QY 407 TCCGTCATATTTTATTTTCAATATGATATAGTATTCATTTCCGCGGAGATTAATC 466
 Db 301 TCCGTCATATTTTATTTTCAATATGATATAGTATTTCCGCGGAGATTAATC 360
 QY 467 GATCGTACCAAGAGACACAGCCATTAAACATCAACCCCAAAATCTCTTGGGG 526
 Db 361 GATCGTACCAAGAGACACAGCCATTAAACATCAACCCCAAAATCTCTTGGGG 420
 QY 527 GCTAAGATTCTCTGTGTGATCTGGGCAATGATGTTCTTACTCTTGTGCTAATCAT 586
 Db 421 GCTAAGATTCTCTGTGTGATCTGGGCAATGATGTTCTTACTCTTGTGCTAATCAT 480
 QY 587 ATTCTGACCAAGAGAGCGAGAGACAAAGATGGAAGAAATGCTTTCCTTAATCA 646
 Db 481 ATTCTGACCAAGAGAGCGAGAGACAAAGATGGAAGAAATGCTTTCCTTAATCA 540

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 11, 2002, 03:17:19 ; Search time 75 Seconds

(Without alignments)
4853.671 Million cell updates/sec

Title: US-09-835-922-1

Perfect score: 1187

Sequence: 1 ggtaccggggatcctctag.....tgtgtcagaactcgttaa 1187

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Database : Issued_Patents_MN:*

- 1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
- 2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
- 3: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
- 4: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
- 5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq:*
- 6: /cgn2_6/prodata/1/ina/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1160	97.7	2247	3	US-08-852-824-1
2	305.6	25.7	1124	4	US-09-221-456-1
3	305.6	25.7	1124	4	US-09-358-740-1
4	302.4	25.5	1488	2	US-08-812-871-2
5	295.4	24.9	1474	2	US-08-467-948A-5
6	295.4	24.9	1474	3	US-08-467-947A-5
7	279	23.5	302	4	US-08-905-223-94
8	250.8	21.1	1027	4	US-09-303-824A-1
9	207.4	17.5	1444	3	US-08-988-876-4
10	129	10.9	1243	1	US-08-702-344-27
11	77.2	6.5	1143	1	US-08-467-125-1
12	77.2	6.5	1143	2	US-08-911-320A-1
13	77.2	6.5	1143	3	US-09-217-101-1
14	71.4	6.4	414	3	US-08-513-874B-353
15	71.4	6.0	357	3	US-08-513-974B-51
16	63.4	5.3	1102	2	US-08-742-440A-5
17	63.4	5.3	1224	2	US-08-742-440A-4
18	62	5.2	1572	1	US-08-041-219A-5
19	62	5.2	1572	1	US-08-417-122-5
20	61.8	5.1	1080	4	US-09-050-159-133
21	60.2	5.1	1080	3	US-08-602-809-3
22	60.2	5.1	1080	3	PCT-US95-16472-3
23	56.8	4.8	1155	3	US-09-053-866-3
24	56.8	4.8	1155	4	US-09-479-130-3
25	52.6	4.4	1224	2	US-08-742-440A-1
26	52.6	4.4	1235	1	US-08-696-770-1
27	52.6	4.4	1235	2	US-09-015-557-1

28	50	4.2	1083	4	US-09-116-498-7	Sequence 7, Appl1
29	48.4	4.1	1083	4	US-09-116-498-11	Sequence 11, Appl1
30	47.6	4.0	1301	2	US-08-467-948A-7	Sequence 7, Appl1
31	47.6	4.0	1301	3	US-08-467-947A-7	Sequence 7, Appl1
32	47.4	4.0	1610	5	US-08-889-108-7	Sequence 7, Appl1
33	47.4	4.0	1610	5	PCT-US94-10358-7	Sequence 7, Appl1
34	47.4	4.0	2160	4	US-08-188-275A-1	Sequence 1, Appl1
35	47.4	4.0	2162	4	US-09-351-198-1	Sequence 1, Appl1
36	47.4	4.0	2162	4	US-09-113-426-1	Sequence 1, Appl1
37	45.8	3.9	1317	1	US-08-153-848-45	Sequence 45, Appl1
38	45.8	3.9	1317	4	US-09-289-843A-45	Sequence 45, Appl1
39	45.8	3.9	1317	4	US-09-088-337B-45	Sequence 45, Appl1
40	45.8	3.9	1317	5	PCT-US93-11153-45	Sequence 45, Appl1
41	45.8	3.9	1664	4	US-09-582-224A-5	Sequence 5, Appl1
42	45.8	3.9	1679	4	US-09-517-605-14	Sequence 14, Appl1
43	45.8	3.9	1737	1	US-08-202-056-4	Sequence 4, Appl1
44	45.8	3.9	1737	1	US-08-076-093A-3	Sequence 3, Appl1
45	45.8	3.9	1737	1	US-08-701-263-3	Sequence 3, Appl1

ALIGNMENTS

RESULT 1									
US-08-852-824-1									
; Sequence 1, Application US/08852824C									
; Patent NO. 6060272									
; GENERAL INFORMATION:									
; APPLICANT: Li et al.									
; TITLE OF INVENTION: Human G-Protein Coupled Receptors									
; FILE REFERENCE: 1488.1220000									
; CURRENT APPLICATION NUMBER: US/08/852,824C									
; CURRENT FILING DATE: 1997-05-04									
; NUMBER OF SEQ. ID NOS: 18									
; SOFTWARE: PatentIn Ver. 2.0									
; SEQ ID NO 1									
; LENGTH: 2247									
; TYPE: DNA									
; ORGANISM: genomic									
; FEATURE:									
; NAME/KEY: CDS									
; LOCATION: (226)..(1251)									
US-08-852-824-1									
Query Match									
Best Local Similarity 97.7%; Score 1160; DB 3; Length 2247;									
Matches 1160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	28	CTGCAGGCTGCATTA	CTACTACTGATGATCA	TTCAAACTCCAGAA	TCACAGTT	87			
DB	147	CTGCAGGCTGCATTA	CTACTACTGATGATCA	TTCAAACTCCAGAA	TCACAGTT	206			
QY	88	ATCAGGTACCAACAA	AGAAATGCAAGCCGTG	CAATTCACCTCTG	CGCTGGGAAC	147			
DB	207	ATCAGGTACCAACAA	AGAAATGCAAGCCGTG	CAATTCACCTCTG	CGCTGGGAAC	266			
QY	148	CAGTCTGACACAG	AGACTCAAAATCACC	AGGCTCTTCCAC	TGCTTACACTGT	207			
DB	267	CAGTCTGACACAG	AGACTCAAAATCACC	AGGCTCTTCCAC	TGCTTACACTGT	326			
QY	208	CCGTGTTTGTG	ACTTATCACAATG	GCCTGCGATG	AGATTTCTTCAATCG	267			
DB	327	CCGTGTTTGTG	ACTTATCACAATG	GCCTGCGATG	AGATTTCTTCAATCG	386			
QY	268	GAGTAAATCAAC	TTTATTTCTTAA	ACACATCTCAT	TTTGTGATCTTCATGAT	327			
DB	387	GAGTAAATCAAC	TTTATTTCTTAA	ACACATCTCAT	TTTGTGATCTTCATGAT	446			
QY	328	TCTGACTTTTC	CAATTCAAATTTCT	AGATGACCAAA	TGGAGACAGACCTGAGAAC	387			
DB	447	TCTGACTTTTC	CAATTCAAATTTCT	AGATGACCAAA	TGGAGACAGACCTGAGAAC	506			
QY	388	TTTGTGTGTC	AAATTTACCTCG	TCATATTTTAT	TTCACAAATGATATAC	447			

|||||
Db 507 TTTTGTGTGCAAGTTACCTCCGTCATATTTTTCACATGTATATGATTTTCATT 566
Qy 448 CTTGGAGCTGATATCATATCGTATCCAGAAAGACACAGCCATTTTAAACATCCAA 507
Db 567 CTGGAGCTGATATCATATCGTATCCAGAAAGACACAGCCATTTTAAACATCCAA 626
Qy 508 CCCCCAAAATCTCTGGGGGCTAAGATTCCTCTCTGTTGTATCTGGGCAATTCATCTT 567
Db 627 CCCCCAAAATCTCTGGGGGCTAAGATTCCTCTCTGTTGTATCTGGGCAATTCATCTT 686
Qy 568 ACTCTCTTCTGCTATGATGATTTGACCAACAGCAGCAGCAGCAAGATGTAAGAA 627
Db 687 ACTCTCTTCTGCTATGATGATTTGACCAACAGCAGCAGCAGCAAGATGTAAGAA 746
Qy 628 ATGCTCTTCTGCTATGATGATTTGACCAACAGCAGCAGCAGCAAGATGTAAGAA 687
Db 747 ATGCTCTTCTGCTATGATGATTTGACCAACAGCAGCAGCAGCAAGATGTAAGAA 806
Qy 688 TCAAGCATTTTCTGATTAATTTCTTAATTTGTTATTTGTTATTTGTTATTTGTTATTT 747
Db 807 TCAAGCATTTTCTGATTAATTTCTTAATTTGTTATTTGTTATTTGTTATTTGTTATTT 866
Qy 748 AGAAGCTATACCGGTATACGTAAGAAAGAGGGGTAGTAAGTCCCGAGAAAAAGT 807
Db 867 AGAAGCTATACCGGTATACGTAAGAAAGAGGGGTAGTAAGTCCCGAGAAAAAGT 926
Qy 808 GAAAGCTAAAGTTTATATATATATATATATATATATATATATATATATATATATAT 867
Db 927 GAAAGCTAAAGTTTATATATATATATATATATATATATATATATATATATATATAT 986
Qy 868 TGCCCGAATTCCTTACACCTGAGCCAAACCGGGGATGCTTGTACATGCTCTGAAAA 927
Db 987 TGCCCGAATTCCTTACACCTGAGCCAAACCGGGGATGCTTGTACATGCTCTGAAAA 1046
Qy 928 TACTGTCTATGTAAAGAGAGACACTGTGTGTTAACTTCTTAAATCATGCTCTGA 987
Db 1047 TACTGTCTATGTAAAGAGAGACACTGTGTGTTAACTTCTTAAATCATGCTCTGA 1106
Qy 988 TCCGTTATCTATTTTCTCTTTCGCAAGTCCCTTCAGAAATTCCTGTATTAAGTCTGAA 1047
Db 1107 TCCGTTATCTATTTTCTCTTTCGCAAGTCCCTTCAGAAATTCCTGTATTAAGTCTGAA 1166
Qy 1048 GTGCCCCAATTCGCAACATCTGTGCCCCAGCAATAGAAAAAGAAAGAGATGTTG 1107
Db 1167 GTGCCCCAATTCGCAACATCTGTGCCCCAGCAATAGAAAAAGAAAGAGATGTTG 1226
Qy 1108 TGACCCAAATGAGAGACTCCATGTAAACAATTAATTAAGAAATATTTCAATCTCT 1167
Db 1227 TGACCCAAATGAGAGACTCCATGTAAACAATTAATTAAGAAATATTTCAATCTCT 1286
Qy 1168 TGTGTTCAAGACTCGTTAAA 1187
Db 1287 TGTGTTCAAGACTCGTTAAA 1306

RESULT 2
US-09-221-456-1
Sequence 1, Application US/09221456
Patent No. 6162899
GENERAL INFORMATION:
APPLICANT: SATHE, GANESH
APPLICANT: HALSEY, WENDY
APPLICANT: MOIR, ALISON
APPLICANT: CHAMBERS, JON
APPLICANT: SEKERES, PHILIP
TITLE OF INVENTION: METHODS OF SCREENING FOR AGONISTS
TITLE OF INVENTION: AND ANTAGONISTS OF THE HNEA81 RECEPTOR
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ratner & Prestia
STREET: P.O. Box 980
CITY: Valley Forge

STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/221,456
FILING DATE: 28-DEC-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/956,975
FILING DATE: 23-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Prestia, Paul F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-70318-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0700
TELEX: 846169
INFORMATION FOR SEQ. ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-221-456-1

Query Match 25.7%; Score 305.6; DB 4; Length 1124;
Best Local Similarity 57.7%; Pred. No. 8.4e-80;
Matches 545; Conservative 0; Mismatches 399; Indels 0; Gaps 0;

Qy 152 CTGTGACCCAGAACATCAAAATACCCAGGCTCTTCCCATGCTCTACACTGCTG 211
Db 137 CGGTGCCCCAGAGACACGTCGATAGTACAGCTGTATTCACACCCCTTACACAGTG 196
Qy 212 TTTTGTGTGAGCTATATCAAAATGCGCGATGAGATTTTCTTCAAAATCCGAGT 271
Db 197 TTCTTACCGGCATCTCTGCTGATTTTGTGCTCTGTGGTGTGTTTGTTCATGCTCCAG 256
Qy 272 AAATCAAACTTTATTTTCTTAAAGAACACAGTATTTCTGATCTCTCATGATTTCT 331
Db 257 TCCCTCACCTTCAATCATCTACCTCAAAAACACTTGTGTGCGCACTGTATTAAGACATC 316
Qy 332 ACTTTCACATCAAAATCTTGTAGTATGCAAACTGGGAACAGGACACAGCACTTTT 391
Db 317 ATGCTTCCTTCAAAATCCCTCTCTGACTCACACCTGGACCCCTGACGCTCAGAGCTTT 376
Qy 392 GTGTGTCAAGTTACCTCCGTCATATTTTATTTTCAATGTATATTTTCAATTTCTG 451
Db 377 GTGTGTGTTTTTCTCGGTGATTTTATGAGACATGATGTGGGATCGTGTGTTA 436
Qy 452 GCACTGATATATATGATGCTACCAAGAAAGACCAAGCCATTTTAAACATCAACCC 511
Db 437 GGGCTATAGCCCTTTCAGAGATTCCTCAAGATCATCAGACCTTTGAGAAATATTTTCTA 496
Qy 512 AAAAATCTTGGGGGCTAAGATTCCTCTGTTGTATGATGAGCAATTCATGTTCTTAC 571
Db 497 AAAAATCTTGGGGGCTAAGATTCCTCTGTTGTATGATGAGCAATTCATGTTCTTAC 556
Qy 572 TCTTTCCTTACATGATTTGACCAACAGCAGCAGCAGCAGCAAGATGTAAGAAATG 631
Db 557 TCCCTGCCAAATATGATTTGACCAACAGCAGCAGCAGCAGCAAGATGTAAGAAATG 616
Qy 632 TCTTTCCTTAAATCAGAGTTGCTAGTGTGCAATTAATTAATTAATTAATTAATTA 691
Db 617 GCTTCTTAAAGGGGCTCTGGGGCTGAATGATGATGATGATGATGATGATGATGATG 676
Qy 692 GTCATTTTCTGATTAATTTCTTAATTTGATTTGATTTGATTTGATTTGATTTGATTT 751


```

SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/812,871
FILING DATE: Filed Herewith
Prior Application DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0237 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1488 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
LIBRARY: MMLR3DT01
CLONE: 568987
US-08-812-871-2

```

```

Query Match      25.5%; Score 302.4; DB 2; Length 1488;
Best Local Similarity 57.5%; Pred. No. 8.4e-79;
Matches 543; Conservative 0; Mismatches 401; Indels 0; Gaps 0;

```

```

QY 152 CTGTGACGACGAGACATCAAAATCACCAGGCTCTCTCCAGCTGCTACAGTCTGCTG 211
DB 131 CGGTGCCCGACGACACCTGGATGATGACAGCTGATTCACAGCCCTCTACAGAGTGT 190
QY 212 TTTTGTGACTTTCACAAATGCGCTGCGATGAGATTTTCTTCAATCCGAGT 271
DB 191 TTCTTGACGGGACCTCGTGATGATCTTGCGCTGTGGTGTGTTGTTCATCCAGC 250
QY 272 AATTCAACTTATTTATTTTCTTAAAGACACAGTCAATTTGATCTTCTCAGATTCG 331
DB 251 TCCCTCACCTTCATCTACCTCAAAACACTTGGTGCCGATGATTAAGACACCTC 310
QY 332 ACTTTCACATCAAAATTTCTAGTATGCCAAATCGGAGACAGACACCTGGAACCTTT 391
DB 311 ATGCTTCCTTCAAAATCTCTCTGACTACACCTGGACACCTGCGACAGCTCAAGCTTT 370
QY 392 GTGTGCAAGTACCTCCGTCATATTTTATTTTCAATGATATGATGATTCATTCCTG 451
DB 371 GTGTGCGTTTCTCTGCGATATTTTAAGACCATGATGTGGCATCGCTGTGA 430
QY 452 GACGTATATGATGATGATGCTACCAAGAACACCAAGCCATTTAAACATCCAAACC 511
DB 431 GGGCTCATAGCCCTTTCAGATTCCTCAAGATCATCAGACCTTGAATATTTTCTA 490
QY 512 AAAAATCTTGGGGCTAAGATTTCTCTGTTGTCATCGGGCATTCATTCATCTACCTC 571
DB 491 AAAAATCTTGGGGCTAAGATTTCTCTGTTGTCATCGGGCATTCATTCATCTACCTC 550
QY 572 TCTTTGCTTGAATGATTTCTGACCAAGGACGACCGACAGACAAAGATGTGAAGAAATG 631
DB 551 TCCCTGCCAATATGATCTTGAACCAAGAAAGAACCAACCATCGCTGTGAAAGTGT 610
QY 632 TCTTTCCTTAAATCAGAGTTCGCTGATGCTGCGATGAATATGATTAATTCATCTGCA 691
DB 611 GCTTTCCTTAAAGGGCTCTGCGGCTGAAGTGAATGCTCAATATGATTAATGATCAAG 670
QY 692 GTCATTTTCTGATTAATTTCTTAATTTGATTTATGATTAATGATTAATGATTAAGAA 751
DB 671 TTTATTTTCTGACCTTTTAAATCTTAATGCTTGTGTTTATGCTGTTATGCAAAAAA 730
QY 752 CTGTACCGGTATACGTAAGAACGAGGGGTGATGATTAAGTCCCAAGAAAGAGTGAAC 811
DB 731 GATATGATTTCTTAATGAAGAGTCAAAATGTAAGACAGAAAAAACAACAAAGAGTGA 790

```

```

QY 812 GTCAAACTTTCATTAATCATGCTGATTCATTTATTTGTTGTTCCATTTGGCC 871
DB 791 GGCAAGTATTTGTCGCGGCGCTCTCTCTTTGTTGCTTTCATTTGCTGCTTTC 850
QY 872 CGAATTCCTTACACCTGACCAACCCGGGATGCTTTGATGCTGCTGTAATACT 931
DB 851 AGAGTTCATATCTCACTCAACCCCAACCAAGACTGATGATGCTCAAAATCAA 910
QY 932 CTGTTCATGTAAGAGACACTGCTGTTGATCTTCTTAATGATGATGCTGATCCG 991
DB 911 CTGTTATTTCTTAAGAAACACTCTTTTGGACCAACTAATTTGATGATCC 970
QY 992 TTCAATATTTTCTTTCGCAAGTCCCTCAAGAAATCTTGAATGATGATGATGCT 1051
DB 971 TTAATATTCATATTTCTATGTAATAAATTCACAGAAAGCTACATGATGATGAG 1030
QY 1052 CCCAATTCGCAACATCTGCTCCAGGACATAGGAAAAAGA 1095
DB 1031 AAGACCACAGCATCAAGCCAAAGAAATCATGACATGCTACAGAGA 1074

```

RESULT 5

```

US-08-467-948A-5
Sequence 5, Application US/08467948A
Patent No. 5998164

```

```

GENERAL INFORMATION:
APPLICANT: LI, YI
APPLICANT: CAO, LIANG
APPLICANT: NI, JIAN
APPLICANT: GENTZ, REINER
APPLICANT: BULT, CAROL J.
APPLICANT: SUTTON III, GRANGER G.
APPLICANT: ROSEN, CRAIG A.
TITLE OF INVENTION: Polynucleotides Encoding Human G-protein
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVE., NW, SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,948A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
Prior Application DATA:
APPLICATION NUMBER: PCT/US95/04079
FILING DATE: 30-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: STEFFE, ERIC K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.1140003/EKS/KLM
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1474 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 62..940

```



```

Db 336 CTTTGTGTCGCTTTTCTTGGTGATATTTATGAGACCATGTATGTGGCATGTCG 395
Qy 447 TCTGGAGATGATTAAGTATGATGCTACCGAAGACCCAGGCCATTTAAACATCA 506
Db 396 TGTTAGGGCATAGGCTTGTGACAGATTCCTCAAGATCATACACCTTTGAGAAATATT 455
Qy 507 ACCCCAAAATCTCTGGGGGCTAAGATTCCTGTGTGATCTGGGCAATTCATGTTCT 566
Db 456 TTCTAAAAAACTGTTGGGGAACCGGTCTCAATCTTCATCTGGTCTTTGGTTC 515
Qy 567 TACTCTCTTGGCTTAACATGATCTGACCAACGAGCGAGCCGAGACAAAGATGTGAGA 626
Db 516 TCATCTCCCGCAAAATGATATCTGTAGCAACAGAGAACACACATCTGTGTGAAA 575
Qy 627 AATGCTCTTCTTAATATCAGAGTGGTCTAGTGGCAATGTAATGTAATTTACATCT 666
Db 576 AGTGTCTCTTAAAGGGGCTCTGTGGGCTGAAGATGGCTCAAAATGTAATATAT 635
Qy 687 GTCAAGTCATTTCTGGATTAATTTCTTAATGTGTAATGTATGATACATTCATCA 746
Db 636 GCCAGTTTATTTCTGACTGTTTATCTTAATGCTTGTTGTTATGTGTTATGCAA 695
Qy 747 AAGAGTGTACCGGTATAGTAAAGAGGAGGTAGTAAAGTCCCGAGGAAAAAG 806
Db 696 AA-AGTATATGATTTCTTATAGAAAGTCCAAAGTAAAGGACAGAAAAACAAAG 754
Qy 807 TGAACGCTAAAGTTTCATTAATCATGCTGTAATCTTTATTTGTTGTTCCATT 866
Db 755 TGGAAAGCAAGATTTGTTGTTGCTGCTGCTCTTCTTGTGTTGCTCCATTTTCAT 814
Qy 867 TTTCCCGAATTCCTTACCCCTGAGCCAAACCCGGATGCTTGTGCTGACGCTGAAA 926
Db 815 TCCCGAGAGTTCATTAATACAGTCAACCAACAAATAGTCACTGTAAGCTGCAA 874
Qy 927 ATACTCTGTTCTTGAAGAAGAGACCTGTGTTAACTTCTTAAATGATGCTG 986
Db 875 ATCACTGTTTATGCTAAAGAAACAACTCTCTTTTGGAGACATAATTTGATG 934
Qy 987 ATCCGTTCACTATTTTCTCTTGTGCAAGTCTTCAAGAAATCTTGTGATGATGCTGA 1046
Db 935 ATCCCTTAATATCATATTTCTTATGTAAAAAATTCACAGAAAGTACCATGTATGCAAG 994
Qy 1047 AGTGGCCCAATCTGCAATCTCTGCCAGACAAATGGAAGAAAAGA 1095
Db 995 GCGAAGAACCCACGATCAAGCAAGAAATCATAGCACTGACAGACA 1043

```

RESULT 7
US-08-905-223-94
Sequence 94, Application US/08905223
Patent No. 6222029
GENERAL INFORMATION:
APPLICANT: Edwards, Jean-Baptiste D.
APPLICANT: Duclert, Aymeric
APPLICANT: Lacroix, Bruno
TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS
NUMBER OF SEQUENCES: 503
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 501 West Broadway
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-3505
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Win95
SOFTWARE: Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/905,223
FILING DATE:
CLASSIFICATION: 536

```

ATTORNEY/AGENT INFORMATION:
NAME: Israelien, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 94:
SEQUENCE CHARACTERISTICS:
LENGTH: 302 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: DOUBLE
TOPOLOGY: LINEAR
MOLECULE TYPE: CDNA
ORIGINAL SOURCE:
ORGANISM: Homo Sapiens
TISSUE TYPE: Brain
FEATURE:
NAME/KEY: s19_deptide
LOCATION: 99..236
IDENTIFICATION METHOD: Von Heijne matrix
OTHER INFORMATION: score 5.4
OTHER INFORMATION: seq VLEFVGLITNGLA/MR
US-08-905-223-94

Query Match      23.5%; Score 279; DB 4; Length 302;
Best Local Similarity 99.3%; Pred. No. 2,6e-72;
Matches 279; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 28 CTGACGGCTGCATTAATCTACTACTGATGATACATTCAAACCTCCAGAAATCAAGTT 87
Db 20 CTGACGGCTGCATTAATCTACTACTGATGATACATTCAAACCTCCAGAAATCAAGTT 79
Qy 88 ATCAGTAAACCAAGAAATGCAACCCGTCGACATCTACCTGCGCTGGAAAC 147
Db 80 ATCAGTAAACCAAGAAATGCAACCCGTCGACATCTACCTGCGCTGGAAAC 139
Qy 148 CAGTCTGTGACACGAGACTACAAATACCCAGGTCTCTTCCACTGCTTACACTGT 207
Db 140 CAGTCTGTGACACGAGACTACAAATACCCAGGTCTCTTCCACTGCTTACACTGT 199
Qy 208 CCGTCTTTTGTGGAATATCAAAATGGCCGAGATGAGATTTCTTAAATCCG 267
Db 200 CCGTCTTTTGTGGAATATCAAAATGGCCGAGATGAGATTTCTTAAATCCG 259
Qy 268 GAGTAATCAACTTATTTATTTCTTAAAGACACAGTCA 308
Db 260 GAGTAATCAACTTATTTATTTCTTAAAGACACAGTSA 300

```

RESULT 8
US-09-303-524A-1
Sequence 1, Application US/09303524A
Patent No. 6238873
GENERAL INFORMATION:
APPLICANT: CHAMBERS, JONATHAN K.
APPLICANT: STEWART, BRIAN R.
APPLICANT: AMES, ROBERT S.
APPLICANT: SARAU, HENRY M.
APPLICANT: FOLEY, JIM
APPLICANT: ARNOLD, ANNE ROMANIC
TITLE OF INVENTION: METHODS OF SCREENING FOR AGONISTS AND
TITLE OF INVENTION: ANTAGONISTS OF THE INTERACTION BETWEEN THE HUMAN KIAA0001
TITLE OF INVENTION: RECEPTOR AND LIGANDS THEREOF
FILE REFERENCE: G950007
CURRENT APPLICATION NUMBER: US/09/303,524A
PRIOR FILING DATE: 1999-04-30
PRIOR APPLICATION NUMBER: 60/083,957
PRIOR FILING DATE: 1998-05-01
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1

LENGTH: 1027
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-303-524A-1

Query Match 21.1%; Score 250.8; DB 4; Length 1027;
 Best Local Similarity 55.4%; Pred. No. 8.9e-64;
 Matches 486; Conservative 0; Mismatches 392; Indels 0; Gaps 0;

155 TGCACGAGAGACACAAATACACCAGGTCCTTCACACAGCTCTACACTGCTGTTT 214
 43 TGCCTCAGAACCTCCGATACACAGATCATCTCTGCTGCTATGCTGCTC 102
 215 TTGTGGACTTATACAAATGCGCTGGCAGTATTTCTTCAATCCGAGTAA 274
 103 ATTGCAGGAATCCACTCAATGAGATGATGATGATGATGATGATGATGAT 162
 275 TCAAACTTTATTTTCTTAAGAACAGATCTTCTGATCTCTCATATTTGACT 334
 163 AAGATTTCATCATCTATCTCAGAACATGTTTGTCTGATCTGATGATGATGAT 222
 335 TTTCATCAAAATTTCTAGATGATGATGATGATGATGATGATGATGATGAT 394
 223 TTTCCTTCAAGATCTCTGATGATGATGATGATGATGATGATGATGATGAT 282
 395 TGTCAAGTACCTCCGCTCAATTTTATTTCAATGATATGATGATGATGATGAT 454
 283 TGCAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 342
 455 CTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 514
 343 CTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 402
 515 AATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 574
 403 TCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 462
 575 TTGCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 634
 463 GTTCAATATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 522
 635 TTCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 694
 523 GAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 582
 695 AATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 754
 583 ATCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 642
 755 TACCGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 814
 643 TTTAAGTCCACCTTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 702
 815 AAATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 874
 703 AATCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 762
 875 AATCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 934
 763 ATCCCTTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 822
 935 TTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 994
 823 CGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 882
 995 ATCTATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1032
 883 ATTATTTCTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 920

RESULT 9
 US-08-988-876-4
 ; Sequence 4, Application US/08988876

Patent No. 6063596
 GENERAL INFORMATION:
 APPLICANT: Lal, Preeti
 APPLICANT: Bandman, Olga
 APPLICANT: Hillman, Jennifer L.
 APPLICANT: Yue, Henry
 TITLE OF INVENTION: G PROTEIN COUPLED RECEPTORS ASSOCIATED
 TITLE OF INVENTION: WITH IMMUNE RESPONSE
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Incyte Pharmaceuticals, Inc.
 STREET: 3174 Porter Drive
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 OPERATING SYSTEM: IBM Compatible
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/988,876
 FILING DATE: Herewith
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J.
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PF-0441 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650-855-0555
 TELEFAX: 650-845-4166
 TELEX:
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1444 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: PROSTUT09
 CLONE: 1650519
 US-08-988-876-4

Query Match 17.5%; Score 207.4; DB 3; Length 1444;
 Best Local Similarity 53.0%; Pred. No. 5.5e-51;
 Matches 467; Conservative 0; Mismatches 411; Indels 3; Gaps 1;

157 CACGAGACATACAAATACACCAGGTCCTCTCCACTGCTCAACATGTCCTGTTT 216
 420 CACCTTACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 479
 217 TGTGGATTTATCAAAATGCGCTGGCGATGAGATTTCTTCAATCCGAGTAATC 276
 480 GGCAAGATCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 539
 277 AAATTTATTTTCTTAAAGACAGATGATGATGATGATGATGATGATGATGATG 336
 540 CAGCTTATTTATTTCTTAAAGACAGATGATGATGATGATGATGATGATGATGAT 599
 337 TCAATTCAAATTTCTTAAAGACAGATGATGATGATGATGATGATGATGATGATG 396
 600 TCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 659
 397 TCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 456
 660 CAGATACATCTCAGTTGTTTATGCAAAACATGATGATGATGATGATGATGATG 719
 457 GATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 516

```

Db 720 GATACATTGATCGCTATCTGAGGTGGTCAAGCCATTGGGACTCTGCGATGACAG 779
Qy 517 TCTCTTGGGGGCTAGATTCCTCTGTGTGTCATCTGGGCAATGTTCTTACTCTCTT 576
Db 780 CATACCTTACAGAGTTTATCTGTGTGTGTGGTATCATGGCTCTTTTGTCTTT 839
Qy 577 GCTTAACATGATTTCTACCAACAGGCAAGCCAGACAGAGATGTGAAGAAATGCTTTT 636
Db 840 GCCAAACATCATCTGACAAATGGTCAAGCCAGACAGAGAAATATCCATGACGCTCAA 899
Qy 637 CCTTAATCAGATGCTGGTCTAGTCTGCGATGAATAGTAATATACATCTGCAAGTCAT 696
Db 900 ACTTAAAGAGCTTTGGGGGTCAAAATGCGATAGCGAGTACCTATGTGAACAGCTGCTT 959
Qy 697 TTTCTGATTAATTTCTTAATGTTATGTATGTATATACACTATTAACAAAGACTGTA 756
Db 960 GTTGTGGCCGCTGCTGATCTGTATGATCGGATGTACATAGCCATATCCAGGTACATCA 1019
Qy 757 CCGGTACATACGTAGAACAGAGGGGTGATAGTAAAGTCCCGAGAAAAGGTGAACGTCAA 816
Db 1020 CAATTCAGAGGCAATTCATAAGT--CAGTCAAGCCGAAAGCAAAACATTAACAGAG 1076
Qy 817 AGTTTCATATATCATGCTGATATCTTATTTGTTGTCCTTCCATTTGCCGAAT 876
Db 1077 CATCAGGCTGTGTGTGCTGTGTATTTTACCTGCTTCTACATATCCTTGTGCGAAT 1136
Qy 877 TCTTACACCTGAGCCAAACCGGGATGCTTTGACTGCACTGCTGAANAATCTGTGT 936
Db 1137 GCCTTCACTTTAGTACACTTACAGAGGCTTTAAGATGAATGATCGACAAATAATCTATA 1196
Qy 937 CTATGGAAGAGACACTGTGTGTTAACTTCTTAAAGCATGGCTGATCCGTTTCAAT 996
Db 1197 TTACTGCAAGAAATTAACATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1256
Qy 997 CATTTTTCCTTGGCAAGTCTTCAAGAAATTCCTGATA 1037
Db 1257 TTACTTTTCACTGTAGTCAATTTTCAAGATGGCTGTCA 1297

RESULT 10
US-08-702-344-27
; Sequence 27, Application US/08702344
; Patent No. 5723315
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John
; APPLICANT: Lavallee, Edward
; APPLICANT: Racie, Lisa
; APPLICANT: Merberg, David
; APPLICANT: Spaulding, Vikki
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; TITLE OF INVENTION: ENCODING THEM
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702,344
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724

```

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1243 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-702-344-27

Query Match      10.9%; Score 129; DB 1; Length 1243;
Best Local Similarity 48.4%; Pred. No. 4,6e-28;
Matches 428; Conservative 0; Mismatches 445; Indels 12; Gaps 2;

Qy 142 GAACACAGCTGCTGACACAGACATCAAAATACCCAGTCTCTCCACTGCTCA 201
Db 201 GACAAACAGTCTGCTCTCTGCCCCAGTTTAAAGATCTGGAGCCATTCACGATTTT 260
Qy 202 CACTGCTGCTTTTGTGAGACTATACAAATGGCTGGAGTGAATTTCTTTCA 261
Db 261 TTAATTAGTTTCTGTTGGAAATTAATGGAATTTTGTGCAAGCTGGCTTTTACA 320
Qy 262 AATCCGAGTAATCAACT--TTATTTATTTCTTAAGACACAGCATTTCTGATCT 318
Db 321 GAAGAAATAGCAATCACAGGTGTGTGACATCTACTTAATTAATTGCTTAACACCGATT 380
Qy 319 TCTCATGATTCGACTTTCCATTCATAAATCTTACGTAGTGGCAAACTGGGAACAGAC 378
Db 381 CTGCTTACTCTGCAATTCACAGTGAANAATGTTGTGACTGGGTGGCCACTTGGAA 440
Qy 379 ACTGAAACTTTTGTGTGACATCTCCGTCATATTTTATTTCAATATATATCA 438
Db 441 GCTGAAGATATTCACAGTGAACAGTACAGCTGCTCATCTATATCAATATGATTTATC 500
Qy 439 TATTTATTCCTGGAGCTAATACATCTGATGCTTACAGACACACAGCCATTAA 498
Db 501 AATTATCTTTAGCAATTTGTGACATTTGACCGCTGCTTACACTGACACACAGCTGCA 560
Qy 499 AACATCCAAACCCCAAAATCTCTGGGGCTAAGTATCTCTGTTGATCTGGGCAAT 558
Db 561 GATCTACCGAAATCAAGAACCCGATTTGCCAAATATGATATCAACGTTGTGTGCTAAT 620
Qy 559 CATGTTCTTACTCTCTTGGCTAATCATGATTTCTGACCAACAGGACCGAGACAGAA 618
Db 621 GGTCTTCTTAATATGATGCAAAATATGATTTCCATCAAAAGATCAAGGAAAGTC 680
Qy 619 TGTGAAGAAATGCTCTTCTTAAATCAGAGTGGCTAGTGTGCAATGAATATGATA 678
Db 681 AATGTGGGTGTATGAGTTTAAAGAAATTTGGAAGAAATTTGGCATTTGCTGACAA 740
Qy 679 TTACATCTGCAAGTATTTCTGATTAATTTCTTAATTTGATTTATGATATACAT 738
Db 741 TTTCAATATGTAGCAATATTTTAAATTTCTAGCCATCATTTTAATATCAATTTGCT 800
Qy 739 CATTACAAAGAACTGACCGGTCATACGTACAGACAGAGGGGTAGTAAGTCCCAAG 798
Db 801 TGTAAATTCGACACTCTACAG-----AAACAAAGATTAAGAAATTAACCAAAATG 851
Qy 799 GAAAAAGGTGACAGTCAAGATTTCATTTATCATGCTGTAATTTCTTAATTTGTTTCC 858
Db 852 GAAAAAGGTCTCTCATCAATCTTTAGTGAACACAGGGCTATCATATATGCTTTTCC 911
Qy 859 TTTCATTTTGGCCGAATCTTACACCCGAGCAACACCCGGAGTGTCTTACAGCAC 918
Db 912 TTACCATATTTGCGAATCCGTAATACCTCAGCCAGACAGAAAGTATATACATGATGCT 971
Qy 919 TGTGAAATATCTGCTGTATGTGAAGAGACACTGTGTGTTACTTCTTAATATC 978
Db 972 AACCAAGATTTACCTCTCAAAAGCAAGAGGCTACAGTCTGCTGCTGTGTGCAACCT 1031
Qy 979 ATGCTGTGATCGTCAATATTTTCTTGTGCAAGTCCCTCAG 1023

```

Db 1032 GTGCTTTGATCTATCTGCTGATCTACCTGCTCAAAAGCATTCG 1076

RESULT 11

US-08-467-125-1

Sequence 1, Application US/08467125
Patent No. 5686597

GENERAL INFORMATION:

APPLICANT: Coleman, Roger
APPLICANT: Au-Young, Janice
APPLICANT: Bandman, Olga
APPLICANT: Selikamer, Jeffrey J.
TITLE OF INVENTION: Thrombin Receptor Homolog
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3330 Hillview Avenue
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,125
FILING DATE:
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Luther, Barbara J.
REGISTRATION NUMBER: 33954
REFERENCE/DOCKET NUMBER: PF-0041 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-852-0195

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 1143 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
LIBRARY: Liver
CLONE: 86700
US-08-467-125-1

Query Match 6.5%; Score 77.2; DB 1; Length 1143;

Best Local Similarity 45.1%; Pred. No. 6.4e-13;
Matches 406; Conservative 0; Mismatches 488; Indels 6; Gaps 3;

QY 130 CTCGCGCGCTGGGAACACAGTCTGTGACACGAGACTACAAATCACCAGGCTCTT 189
Db 111 CTCGCAACACCAATGTTACTACTGTCATGATGATGATGATGATGATGATGAT 170
QY 190 CCCACTGCTACACAGTCTGCTGTTTGTGAGCTATACCAATGAGCCGCGGAGTAG 249
Db 171 AACCAATCTCTACTCTGTATTTTTCATGCTGAGTGTGGAAACATTAACGCCCTTA 230
QY 250 GATTTTCTTCAAAATCCGAGTAATATCAACTTATTTTCTTAAAGAACACAGT 306
Db 231 TGTATTTCTGGGTATTCACCTAAAGAAATTCATTCATTAATTAATTAACGTAGC 290
QY 307 CATTTCTGATTTCTCATGATTCGACTTTTCATTAATTTCTTACTGATGCCAACT 366
Db 291 CATTCGACAGCTCTACATCTCTGCTCCCTTCGGAATATGATATTAATTAACCA 350
QY 367 GGAAGACAGACCTGAGAGACTTTTGTGTCAAGTTACCTCCGATATTTTATTTAC 426
Db 351 AAACAGTGGACACTAGTGTGATTTCTGTGCAAGTTGTGGAAACACTGCTTTATATGA 410

QY 427 AATGTAATCAGTATTTTCATCTCTGGAGCTGATATACATATGATGCTACCAAGACAC 486
Db 411 CATGTACATTTAGCATTTATTTTCTGTGGATTCATGATTTGGATTCCTATATAAAATTA 470
QY 487 CAGGCAATTTAAACATCAACCCCAAAATCTCTTGGGGCTAAGATTCCTCTCTGTGT 546
Db 471 TGGTCTATACAGCAACGAGGCAATTAACCAACCAAGTA--TTATGTCTGTGT 528
QY 547 CATCTGGCATTCATGTTCTTACTCTCTTTCCTTACATGATTTCTGACCAACGACACC 606
Db 529 ATAGTATGATGCTGTCTGTGT--GATTCCTACTATGATTTATTAACCTTAACAA 587
QY 607 GAGAGACAGATGTGAAGAAATGCTCTTTCCTTAATCAAGATTCGCTAGTCTGGCA 666
Db 588 AGGAGGCAATATTCACAAATGTGTTTCCATTAACAGATTAAGCATTAACGCAAAAGAGA 647
QY 667 TGAATATGTAATTAATCATCTGTCAAGTCATTTCTGATTAATTTCTTAATTTATTTGT 726
Db 648 AGCCATTTTAACTTCATCTGTGTGATGTTCTGCTTAATTTCTTAATTAATCT 707
QY 727 ATGTTATACACTCATTAACAAAGAACGTCACCGCATACGTAAGAACGAGGCTAGG 786
Db 708 TTCAATATTAAGATTTGGGAATCTATGAGATTTCTAAAGAGGCTCAAAATTTCC 767
QY 787 TAAAGTCCCGAGAAAAGGTGACGTCAAGTTTTCATTAATGATGCTATTTCTTAT 846
Db 768 TAATTTCTGTAATATGACCTACAGCTGTAACCTTTATTTGATTTATCATTTTATTC 827
QY 847 TTGTTTGTCTCTTCATTTTGTCCCGAATTCCTTACACCTGAGCAACCCGGGATGT 906
Db 828 TATATGTGTGGTCTCTTCATGCTTTCGATTCATTCATTTCTTACAGCTTAAGT 887
QY 907 CTTGACATGCACTGCTGAATATCTGTTTATGTAAGAGACACTGCTGTGTATAC 966
Db 888 ATCATCTTCTACTGGAAGAAATTTGTACAAACCATGATGATGATGATGATGATGATG 947
QY 967 TTCCTTAATGCAATGCTGATGCTGATGCTGATGATGATGATGATGATGATGATGAT 1026
Db 948 ATCTTCAATATGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1007

RESULT 12

US-08-911-320A-1

Sequence 1, Application US/08911320A
Patent No. 5869633

GENERAL INFORMATION:

APPLICANT: INCYTE PHARMACEUTICALS, INC.
TITLE OF INVENTION: THROMBIN RECEPTOR HOMOLOG
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/911,320A
FILING DATE: August 14, 1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/467,125
FILING DATE: 6-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0041-1 DIV
TELECOMMUNICATION INFORMATION:

Oy	367	GGGAACAGGACACACTGAGCACTTTGGGAGCTCAAGTTACCTCCGATCTTTTATTTTAC	426
Db	351	AAACAAGTGAACACTAGGTGTGATTTCTGTCAGAGGTTGGGAACACTGCTTTATATGAA	410
Oy	427	AATGTATATCAGTATTTTCATTTCCCTGGGACTGATATACATCGCTACAGAACACC	486
Db	411	CATGTACATTAGCACTTATTTTCCCTGGATTCATCACTTTGGATCCGTATATATAATTA	470
Oy	487	CAGGCCATTTAAACAATCCAAACCCAAATCTCTGGGGGCTAGATTCCTCTGTTGT	546
Db	471	TGGGTCTAATACAGCAAGGAAGCAATMACAACCAACAAAGAA--TTTATGTCTGTGT	528
Oy	547	CATCTGGGCACTCATCTTTCTTACCTCTTTCCCTTAACATGATTTCTGACCAAGACACC	606
Db	529	ATACTATGATGATGCTGTGCTTGGT--GGATTCCTAATCATATGATTTATTTAACTTAAGA	587
Oy	607	GAGAGCAAGAATGTGAAGAAATGCTCTTTAATCATAGAGTTGGTCTGTGCTGGCA	666
Db	588	AGGAGGCCATTAATCCCAATGTGTTCATTCACAGATTAAGCATMACGCAAAAGGAGA	647
Oy	667	TGAATATGTAATATACATCTGTCAAGCATTTTCTGATTAATTTCTTAATGTTATGT	726
Db	648	AGCCATTTTATCTCTACTTCTTGTGGTAAATGTTCTGGCTAATTTTCTTAATATCT	707
Oy	727	ATGTTATACACTCATTTACAAAGAAGCTACCGGTCATACGTAAAGAAGAGGGGTAGG	786
Db	708	TTCATATATTAAATGTGGAAAGATCATATGAGGATTTCTAAAGAGAGCTCAAAATTTCC	767
Oy	787	TAAAGTCCCAAGAAAAGAGTGAACAGTCMAAGTTTCATATTCATTCGTCTTAT	846
Db	768	TAAATCTGGTAATATATGCCACTACAGTACGCTGAACCTCTTATATGATCATATTTTAC	827
Oy	847	TTGTTTGTTCCTTTCCATTTTGGCCGAAATCTTACACCTGAGGCAAAACCCGGGATGT	906
Db	828	TATATGTGTGGTCCCTATCATGCTTTGATTCATCTAATTTCTTTCACACTAAATGT	887
Oy	907	CTTTGACTGCACATCTGTAATAATACCTGTTCTATGTGCAAGAGAGACACTGTGCTTAAC	966
Db	888	ATCATCTGTGTACTGGAAGAAATTTGTCAAAACCAATGAGATCATCTGTGTTCTGC	947
Oy	967	TTCCCTTAATGCAATGCTGTGATCGGTTCAATCTTTTTCCTTTGCAATCCCTCAGANA	1026
Db	948	ATCTTTCATATAGTGTGTATATCCAGTCAGTCATGTTTCCGTGATGTCAGATTAATTCGCA	1007

DB	Query Match	Similarity	6.4%	Score 75.6	DB 3	Length 414
DB	19	GACCGGTAATTTCAACATATGTGAAGCCCTTTCACATGCTCCTCGCGCCGCAACGTC	189	Conservative	0	Mismatches 189; Indels 0; Gaps 0;
DB	527	GCTAAGATTTCTCTGTTGTATCTTGCGCATTCATGTTCTTACTCTCTTTGGCTTAACATG	586			
DB	79	AGCAACACTCGTCCTCGTGGTGTGTTGCTCATGCTCCTCGCGCCGCAACGTC	138			
DB	587	ATTTCAGACCAACAGCAGCCGAGACCAAGATGTGAAGAATGCTTTCCTTAATCA	646			
DB	139	ATTTCACACCAACAGAGATTAGACACGTGACGCGCAGATTAATTCATGAGACTTAAC	198			
DB	647	GAGTTGGTCTAGTCTGGCATGAATTAATTAATTAATTCATCTGTCAAGTCAATTTCTGATT	706			
DB	199	GAACTGGGCGCGCAATGGCGACCAAGCGCTCAACACTCATCTTGTGGGATTTTCTGGCTT	258			

